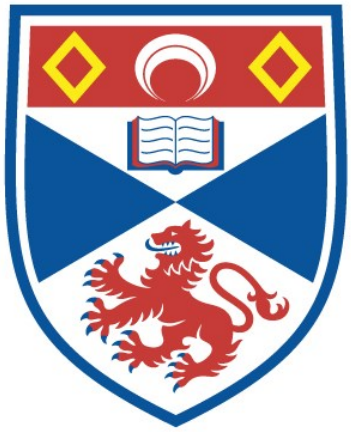
LIVING AND LEARNING TOGETHER: INTEGRATING DEVELOPMENTAL SYSTEMS THEORY, RADICAL EMBODIED COGNITIVE SCIENCE, AND RELATIONAL THINKING IN THE STUDY OF SOCIAL LEARNING

Murillo Pagnotta

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



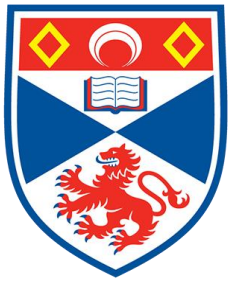
2018

Living and learning together: Integrating developmental systems theory, radical embodied cognitive science, and relational thinking in the study of social learning

Murillo Pagnotta

Living and learning together: Integrating developmental systems theory, radical embodied cognitive science, and relational thinking in the study of social learning

Murillo Pagnotta



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

March 2018

Declarations

Candidate's declaration

I, Murillo Pagnotta, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 80000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in March 2014.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 08/08/2018 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 08/08/2018 Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Murillo Pagnotta, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date 08/08/2018 Signature of candidate

Date 08/08/2018 Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Murillo Pagnotta, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 08/08/2018 Signature of candidate

Contributions

Chapter 6

Prof. Greta Defeyter provided feedback on the study design. Lindsey Stevenson, a staff member of the Life Science Centre at the time of the study, acted as the facilitator in the activity. Liesbeth Verlinden, biologist, double-coded the video data for use in tests of inter-observer reliability. Dr Moreno Coco provided feedback on the interpretation of the results. All other work was carried out by me.

Chapter 7

Dr Moreno Coco provided feedback on the experimental design, assisted setting up the eye tracking system, and provided feedback on the interpretation of the results. John Jennings, a school teacher in Edinburgh, acted as the facilitator in the activity. All other work was carried out by me.

General Acknowledgements

Foremost I thank my supervisor, Kevin Laland, for providing me with the opportunity to join his lab and undertake my PhD in such a vibrant environment. Thank you for your encouragement and support.

I wholeheartedly thank Susan Oyama, Tim Ingold, and Christina Toren, who have been extremely influential in my development over the past few years. I also thank Saadi Lahlou for the invitation to visit the London School of Economics and Moreno Coco for his collaboration in both empirical studies. Thank you to my friend Jerry Hogan for the initial support that made this all possible. I had the chance to visit the Konrad Lorenz Institute for Evolution and Cognition Research in Austria for a period of six months and I gratefully acknowledge their support.

I met many wonderful people in St Andrews along the way, and some deserve being mentioned explicitly. Thank you to James Ounsley and Charlotte Brand, with whom I shared the Mujetty Mansion, as well as Elena Miu, Cara Evans, Camille Troisi, Daniel Cownden, Daniel van der Post, Jessica van der Wal, Joe Stubbersfield, Wataru Toyokawa, and Emily Burdett. On a slightly different note, I also thank Ewa Bieniecka and Philipp Liehm for introducing me to the world of Argentine Tango. You have all made my time in St Andrews unforgettable.

I am indebted to Lindsey Stevenson, the maker-in-residence at the Life Science Centre, and John Jennings, a school teacher and a friend, who have given up their time to assist me in the empirical studies.

I am also extremely grateful to Katherine Meacham for her kindness, support, and efficiency in dealing with administrative matters throughout the years.

Finally, I would like to thank my family and my life partner, Édipo Castro, for all the love and support.

Funding

This work was supported by the John Templeton Foundation (grant SBE1-XOBJ10 awarded to Prof. Kevin Laland), by the University of St Andrews (St Leonard's Fee Scholarship awarded to me), and by the Konrad Lorenz Institute for Evolution and Cognition Research (Writing-up Fellowship awarded to me).

Abstract

Behavioural scientists argue that ‘social learning’ provides the link between biological phenomena and cultural phenomena because of its role in the ‘cultural transmission’ of knowledge among individuals within and across generations. However, leading authors within the social sciences have proposed alternative ways of thinking about social life not founded on the Modern oppositions including nature-culture, biology-culture, body-mind, and individual-society. Similarly, the distinction between a domain of nature and a domain of nurture has also been extensively criticized within biology. Finally, advocates of ‘radical embodied cognitive science’ offer an alternative to the representational-computational view of the mind which supports the conventional notion of culture and cultural information. This thesis attempts to integrate developmental systems theory, radical embodied cognitive science, and relational thinking, with the goal to bring the field of social learning closer to these critical theoretical developments. In Chapter 2, I find no justification for the claim that the genome carries information in the sense of specification of biological form. Chapter 3 presents a view of ontogeny as a historical, relational, constructive and contingent process. Chapter 4 uses the notions of environmental information, abilities, affordances, and intentions to make sense of behaviour and learning. In Chapter 5, I argue that the notion of social learning can be understood in terms of relational histories of development rather than in terms of transmission of information. I then report empirical studies investigating behavioural coordination and social learning consistent with this theoretical framework. Chapter 6 presents evidence that dyads in a joint making activity synchronize their attention constrained by their changing situation and that coordination of attention is predictive of implicit and explicit learning. Chapter 7 presents evidence that joint attention does not require gaze following and that attentional coordination is predictive of learning a manual task. Together, these theoretical and empirical studies suggest a new way of thinking about how humans and other animals live and learn socially, one that is consistent with critical theoretical and philosophical developments that are currently neglected in the literature on social learning.

Contents

Chapter 1 General Introduction

1.1 The idea of a (dis)continuity between nature and culture

One of the epistemological foundations of Modernity (roughly, the historical period after the Middle Ages) has been the belief that the human condition corresponds to a radical separation from the rest of nature, often described as due to culture or society (Descola & Pálsson, 1996; Latour, 1993). The term 'culture' itself has been used within the social sciences in a variety of different contexts and there is no consensual definition (see, for example, Geertz, 1973; Ingold, 2000, 2011a; Kroeber, 1917; Kroeber & Kluckhohn, 1952; Lévi-Strauss, 1966; Tylor, 1871). Nevertheless, it had generally been uncontentious that, whatever it is that makes humans unique, we could say that it belonged to the cultural domain. This is reflected in the historical separation between the natural and social sciences.

However, the idea of a sharp distinction between a domain of nature and a domain of culture has been under criticism in many disciplines. Take the example from paleo-anthropology. A long-standing, influential hypothesis, called "the human creative revolution" model, suggests that symbolic language – and, with it, culture – emerged abruptly in the hominin lineage around 50 thousand years ago (Klein, 2009). There was indeed rich and solid empirical evidence from European excavation sites, as well as cave paintings, supporting this model. Invoking the sudden appearance of the domain of culture as the result of a 'creative revolution' supports and is supported by a worldview that sets humans apart from other animals. Recent paleoanthropological evidence, however, especially from African Middle Stone Age study sites, suggest that such an abrupt passage from a 'state of nature' to a 'state of culture' has never occurred (d'Errico & Stringer, 2011; McBrearty & Brooks, 2000).

The idea of culture as a domain related to, but ultimately distinct from, the biological process that is the living body acting in the world, has also been under criticism within anthropology. Ethnographers describing the immense variation in human ways of living have accumulated accounts for over a century which collectively indicate that thinking in terms of a nature-culture or nature-society opposition is neither a human universal nor a logical imperative, but rather a historical peculiarity within our own, so-called Western, European, Modern tradition (Descola & Pálsson, 1996). French anthropologist Claude Lévi-Strauss, the founder of Structural Anthropology, argued that the separation between nature and culture had had "the force almost of dogma" (Lévi-Strauss, 1992, p. 28) within the discipline. Throughout his work, he argued that culture should be taken as a heuristic tool and a relational concept – i.e., one that does not point to some concrete thing but to a differential relation between two ethnographic groups. In his words (Lévi-Strauss, 1963, p. 295):

> What is called a 'culture' is a fragment of humanity which, from the point of view of the research at hand and of the scale on which the latter is carried out, presents significant discontinuities in relation to the rest of humanity. [. . .] Accordingly, the same set of individuals may be considered to be parts of many different cultural contexts: universal, continental, national, regional, local, etc., as well as familial, occupational, religious, political, etc.

While the notion of culture still thrives in parts of contemporary anthropology, an increasing number of influential authors have taken a more radical step. What they have in common is a dissatisfaction with accepting a set of Modern distinctions as a reasonable starting point for their studies – including between nature-culture, nature-society, biology-culture, body-mind, individual-society (Descola & Pálsson, 1996; Ingold, 2000, 2011a; Ingold & Pálsson, 2013; Latour, 1993; Toren, 2012). These distinctions have been taken for granted in the history of anthropology until recently (and indeed still prevail in most disciplines as well as in common parlance), but now these authors argue for the need to find theoretical alternatives that are not sustained by them. The Modern notion of culture rests on a view of the mind as separate from the world which it must 'represent' in order to perceive it, think about it, and act upon it. Thus, authors critical of the nature-culture or biology-culture divide, such as the influential anthropologists Tim Ingold and Christina Toren, align themselves against the representational-computational approach in

the cognitive sciences. In both authors' work, the life of humans, and indeed of all living beings, is seen as a historical, relational process in which form is continually generated and transformed rather than being encoded (as genetic or cultural information) and expressed (in development or behaviour).

Among others, Tim Ingold has repeatedly, carefully, and passionately, argued that "it makes no sense to speak of 'culture' as an independent body of context-free knowledge" (Ingold, 2000, p. 37) or "as an internalised system of rules and meanings *as distinct* from manifest behaviour patterns and their artefactual products" (p. 159, his emphasis). Instead of taking the idea of a mind (or the individual) separated from the world as our point of departure, as implicit in many notions of culture related to the representational view of the mind, he argues we should take "the human condition to be that of a being immersed from the start, like other creatures, in an active, practical and perceptual engagement with constituents of the dwelt-in world" (Ingold, 2000, p. 43). Thus, human so-called 'cultural' behavioural variation should be understood, not in terms of representations being transmitted between people but as variation in the skills embodied in the process of development of each human organism engaged with his environment. Ingold (2000) acknowledges that his view was profoundly influenced by the work of philosopher of biology Susan Oyama on developmental systems theory as well as James Gibson's ecological approach to visual perception.

Similarly, in her studies of child development and other topics, anthropologist Christina Toren (2008, 2012), makes no reference to culture or mental representations. These terms, along with the associated distinctions including biology-culture and body-mind, do not play the epistemic role of starting points. Instead, her starting point is the idea of social living, including learning, as a micro-historical process. Mind is not conceived of as a function of the brain "but of the whole human being in intersubjective relations with others in the environing world" (Toren, 2012). Toren (2012) acknowledges the influence in her approach of the autopoietic theory of Maturana and Varela (1980), the developmental work of Piaget, and phenomenology.

The notion of a sharp distinction between humans and nonhuman animals in terms of culture has also been under criticism within the natural sciences. Raised in the general atmosphere of Buddhism, biologist Kinji Imanishi, the founder of Japanese primatology,

did not believe in a radical separation between humans and other living things, and argued that other animals might exhibit something like culture. To refer to this possibility, he coined a neologism based on the sound of the English word: "kaluchua" (Nakamura & Nishida, 2006). Imanishi (2002) emphasized the unity and harmony of nature (humans included), the integrated and inseparable system formed by organism and its environment, and the need to recognize the simultaneous existence of similarities and differences among things. With this conviction, he appropriated and adapted methods originally from the social sciences, including identification of individuals and long-term observations of their behaviour and social relations, and applied them to the study of wild horses in Mongolia and, more famously, Japanese monkeys (*Macaca fuscata*) on the Koshima Island and chimpanzee in Africa.

Starting in the sixties with Imanishi, Jane Goodall and others, field studies in Africa revealed that chimpanzees live in social units (or communities) and establish lifelong social relationships (Nishida, 1968); use various tools (Goodall, 1964), including hammers to open hard nuts (Boesch & Boesch, 1981, 1984); hunt cooperatively or collectively (Boesch & Boesch, 1989); and might even actively teach in some circumstances (Boesch, 1991). It became clear that the behaviour of the chimpanzee, our most closely related species, was much more complex, diverse, and humanlike – or, conversely, that human behaviour was much more apelike – than ever realized. Following his ground-breaking research on the variability of chimpanzee behaviour across study sites, William McGrew argued throughout his career that the term 'culture' should apply to chimpanzees as well (McGrew, 1992, 2004; McGrew & Tutin, 1978; Whiten et al., 1999). I have looked at the history of attribution of culture to nonhuman animals and the tensions with anthropology in my masters dissertation (Pagnotta, 2014).

Regardless of the controversies involving the term and ignoring decades of critical debate within the social sciences (Descola & Pálsson, 1996; Ingold & Pálsson, 2013; Latour, 1993), behavioural scientists appropriated the conventional, Modern anthropological notion of culture (such as summarized in Kroeber & Kluckhohn, 1952) and redefined it to avoid any explicit reference to humans, language or symbol use. In one influential paper, Whiten et al. (1999) say that "a cultural behaviour is one that is transmitted repeatedly through social or observational learning to become a population-

level characteristic". Thus redefined, the term could now be used more inclusively such that not only humans, but other species too, could be said to have culture. After decades of research on an increasing list of taxa, there is particularly good evidence of the phenomena described in such terms in chimpanzees (Hobaiter, Poisot, Zuberbuhler, Hoppitt, & Gruber, 2014; McGrew, 1992; Whiten et al., 1999), bonobos (Hohmann & Fruth, 2003), orangutans (van Schaik et al., 2003), whales and dolphins (Allen, Weinrich, Hoppitt, & Rendell, 2013; Whitehead & Rendell, 2015), birds (Aplin et al., 2015; Slagsvold & Wiebe, 2007), and fishes (Helfman & Schultz, 1984; Warner, 1988).

The idea that culture is not restricted to humans is one step away from the Modern view. However, the step consists mainly in arguing that humans are not the only animals to be both biological and cultural beings. There is a common argument in this literature (e.g., McGrew & Tutin, 1978; Whiten et al., 1999) that goes as follows: (1) anthropologists cannot agree on a definition of culture; (2) therefore, biologists are justified in providing their own definition; (3) with culture thus redefined, it is a concept that applies to nonhuman animals; (4) with this terminological issue settled, we can get on with the study of interest. This argument suggests that the use of the term 'culture' in relation to nonhuman animals never intended to question the Modern opposition between two domains (nature-culture) or the related oppositions as a reasonable premise. This was probably the case because behavioural scientists were focusing on the observations accumulating from their empirical work on animal behaviour in the field and in the lab more than on the critical discussion going on in the social sciences about the constitution of Modernity (Latour, 1993) and the establishment of the Modern Sciences (Stengers, 2000).

I find the observations described in the literature on animal learning fascinating. However, instead of assuming as a starting point an inclusive definition of culture and then focusing on the mechanisms responsible for the transmission of cultural information, this project attempts a more radical step. I want to avoid the nature-culture distinction altogether, along with the associated set of problematic distinctions nature-society, individual-society, body-environment, nature-nurture, gene-environment, innate-acquired, biology-culture, and body-mind. My intention in attempting this is to bring studies of 'social learning' in humans and other animals conceptually closer to the

contemporary anthropological criticism to the Modern oppositions (Descola & Pálsson, 1996; Ingold, 2000, 2011a, 2013; Ingold & Pálsson, 2013; Toren, 2012). Thus, the aim of this project is not simply to avoid sensitive terms but to begin the exploration of a new way of thinking about the phenomena of interest.

1.2 Social learning as 'transmission of information'

The notion of 'social learning' is central to this debate because it is considered to provide a conceptual link between biological phenomena and 'cultural' phenomena. This is the case because categories of social learning, such as stimulus/local enhancement, emulation, and imitation, are said to underlie the "cultural transmission" of knowledge and behaviour among individuals within and across generations, thus providing the mechanism that might lead to the establishment of group-typical behaviour patterns (Hoppitt & Laland, 2013; Mesoudi, 2011; Richerson & Boyd, 2005). Much of the currently influential and representative work on social learning and culture (such as Boyd & Richerson, 1985; Heyes, 1994, 2012, 2016; Hoppitt & Laland, 2013; Laland, Odling-Smee, & Feldman, 2001; Mesoudi, 2011; Richerson & Boyd, 2005; Tomasello, 1999; Tomasello, Carpenter, Call, Behne, & Moll, 2005) expresses or, at least partially, endorses a set of inter-related ideas about development, inheritance, evolution, cognition, and 'culture.' These ideas come from different fields including biology, the cognitive sciences, and some approaches to culture theory.

From biology, this includes the ideas that cellular metabolism is fundamentally controlled by information encoded in the genome (Alberts, Wilson, & Hunt, 2008; A. J. F. Griffiths, 2008); that development is mostly controlled by genetic information, possibly complemented by or interacting with information in the animal's environment (Maynard Smith, 2000; Mayr, 2004); that inheritance is about the transmission of genetic information (Futuyma, 1998; Mayr, 2004; Ridley, 2004); and that evolution is about changes in genetic information (Futuyma, 1998; Mayr, 2004; Ridley, 2004). The main idea coming from the cognitive sciences is that the central nervous system is fundamentally a structure that processes information (Fodor, 1975; Thagard, 2014). And the main idea coming from some approaches to cultural theory is that cultural processes are sustained by, and interact with, but are in some strong sense distinct from, biological processes (Kroeber, 1917; Kroeber & Kluckhohn, 1952).

Crucially, behavioural scientists studying social learning and culture added some of their own ideas to the list above. Arguably the main idea is that development – more specifically, the development of behaviour patterns through learning – depends on the 'genetic information' involved in making the body (including the animal's cognitive structure responsible for learning), on 'information' acquired through individual learning, *and also* on 'cultural information' acquired through social learning (Hoppitt & Laland, 2013). As Richerson and Boyd (2005, p. 6) put it, "we are largely what our genes and our culture make us". Another idea is that inheritance includes the transgenerational transmission of 'genetic information' *and also* 'cultural information' (Mesoudi, 2011; Odling-Smee, Laland, & Feldman, 2003; Richerson & Boyd, 2005). As a consequence, evolution is viewed in terms of changes in the 'genetic information' *and also* changes in the 'cultural information' transmitted across generations in the population (Mesoudi, 2011; Odling-Smee et al., 2003; Richerson & Boyd, 2005). With regard to cognition, learning is thought of in terms of 'acquiring information' *either* socially *or* asocially (Heyes, 2012; Hoppitt & Laland, 2013).

It should be evident, given the way I chose to present the set of ideas above, that the term 'information' figures prominently in them. Consistent with this, the notion of culture, as currently used by behavioural scientists, is fundamentally associated with the idea of information being transmitted between individuals and across generations socially *rather than* genetically (Hoppitt & Laland, 2013; Mesoudi, 2011; Richerson & Boyd, 2005). As Hoppitt and Laland (2013, p. 11) say, "culture is built upon socially learned and transmitted information... and does not apply to genetic information or to knowledge and skills that individuals acquire alone". Some authors even identify 'culture' with 'information' and 'mental state': "Culture is information capable of affecting individuals' behavior that they acquire from other members of their species through teaching, imitation, and other forms of social transmission... by information we mean any kind of mental state, conscious or not, that is acquired or modified by social learning and affects behavior" (Richerson & Boyd, 2005, p. 5).

If my characterization is indeed representative, 'information' plays a central epistemic role in the study of social learning and culture. However, it is not at all clear that the term means the same thing when used in reference to different phenomena such

as the molecular processes within cells (metabolism), the typical sequence of changes throughout the lifespan (development), the similarities in the developmental trajectories of organisms related by genealogy (inheritance), the physiological processes within the central nervous system (perception, learning, cognition in general), the organism-environment interaction in real time (behaviour), the establishment, persistence, and change of population-level patterns of behaviour across generations (culture), or mental states. Using the same term without a clear understanding of what it might mean in each case can lead to confusion and equivocation, especially when crossing disciplinary borders. For this reason, I attempted to clarify the different meanings of the term 'information' in relation to these different contexts and I do my best to use the term consistently throughout.

A common thread runs through the set of inter-related ideas about development, inheritance, evolution, cognition, and culture I mentioned above. I will refer to it, following Oyama (1985/2000), as preformationism, understood as a way of thinking about how form arises in nature. Preformationism indicates the idea that form is already present in some configuration before it appears in another configuration. With regard to contemporary biology, this way of thinking supports, and is supported by, the metaphor that "the genome... is a set of instructions on how to make [an organism]" (Maynard Smith, 2000). Thus, biologists conventionally say that some aspects of biological form are already specified by 'information' encoded in the genome before they appear in the phenotype as a result of developmental processes. With regard to the behavioural sciences, this way of thinking supports, and is supported by, the idea that behaviour is the expression of the activity of the nervous system (Hogan, 2017). Thus, behavioural scientists conventionally say that some aspects of behavioural form are already specified by 'information' encoded in the brain before they appear in overt behaviour as the expression of neurophysiological activity. Mesoudi (2011, p. 2) gives an illustrative example of this preformationist way of thinking in the context of social learning and culture:

> Whereas genetic information is stored in sequences of DNA base pairs, culturally transmitted information is stored in the brain as patterns of neural connections (albeit in a way that neuroscientists are only beginning to understand), as well as in extrasomatic

> codes such as written language, binary computer code, and musical notation. And whereas genetic information is expressed as proteins and ultimately physical structures such as limbs and eyes, culturally acquired information is expressed in the form of behavior, speech, artifacts, and institutions.

1.3 This thesis

My intention is to avoid this way of thinking altogether and to begin the development of an alternative. I do this inspired by a set of authors. Susan Oyama (1985/2000, 2000b, 2015) offers a thorough critique of the various versions of the nature-nurture opposition and suggests a processual, relational way of thinking about development, inheritance, and evolution in which history and context-dependency are in the foreground. This way of thinking is known as 'developmental systems theory', where 'theory' is meant in a broad sense, as a general perspective. The term 'developmental system' refers to a target entity (for example, the animal under study) and the changing set of interdependent influences and interactants that make a difference in its persistence and change over time. These interactants might be detected at any scale of size and time and be located on either side of its boundaries (such as the cell membrane or the skin). I will also indicate a developmental system by using expressions such as 'the organism-environment system' or 'the organism-in-the-environment'. Assuming a developmental systems perspective as a starting point involves adopting a view of life as a set of nested, multiscale processes unfolding in time and space; a view of causality in terms of relations of mutual constraints rather than in terms of linear antecedent-consequent relations; and a view of control as emerging in the contingent relations among interactants rather than as localized in a privileged factor (say, the genes in development and mental representations in behaviour). Following this approach, I conceive of ontogeny – the life span of organisms, in all its aspects – as a historical achievement, a process in which biological and behavioural forms are not 'encoded' and 'expressed' but rather continually generated and transformed.

Because behaviour is one aspect of ontogeny, adopting a developmental systems perspective of ontogeny invites, if not in fact requires, exploring an alternative way of thinking about cognition as well. I therefore explore works that have been critical of the dominant representational-computational approach in the cognitive sciences and advocate for ecological, dynamical, enactive views instead, including Chemero (2009);

Fogel (1993); E. J. Gibson (1988); J. J. Gibson (1979/2015); Kelso (1995); Maturana and Varela (1980, 1987); Thelen and Smith (1994); Thompson (2007); Varela, Thompson, and Rosch (1991). It has become common to refer to this body of work as advocating for a "radical embodied cognitive science" (Chemero, 2009). The dominant way of thinking in the cognitive sciences views the cognizing organism as an entity separate from the world, which it must represent in order to perceive it, think about it, and act upon it. The alternative way of thinking, indicated by the label radical embodied cognitive science, starts with a relational view of the organism-in-the-world as a unitary system of causal influences flowing in time, in which both the cognizing organism and its environment are continually being differentiated. In my opinion, this way of thinking about cognition is consistent with the developmental systems perspective. The authors listed above are the main influences in how my own ideas persisted under transformation over the past years.

This project has two complementary components, one conceptual-theoretic (Chapters 2 to 5) and one empirical (Chapters 6 and 7). They are presented as one following the other but in fact they ran somewhat in parallel, mutually influencing each other. In Chapter 2 I begin to explore different meanings of the word information in relation to biological phenomena in an attempt to clarify how the term might be used consistently. I then attempt an original integration of ideas from developmental systems theory, radical embodiment, and relational thinking. My goal is to develop a processual, relational way of thinking about ontogeny (Chapter 3), behaviour (Chapter 4), and social life (Chapter 5) which might be used to explore the phenomena described by the literature on social learning and culture in new ways. I then move on to the empirical component of the project, in which I report two studies that were conceived, carried out, and analysed motivated by this alternative way of thinking. I focused on phenomena that would conventionally be considered cases of 'transmission of cultural information' and viewed them from a processual, relational perspective. Chapter 6 presents the results of a thoroughly exploratory study conducted in a Science Museum with children participating in a joint making task assisted by an instructor. Chapter 7 reports a study conducted with adults in the more controlled setting of an eye tracking laboratory at the University of Edinburgh. Finally, Chapter 8 presents a brief overview of the findings and shortcomings of this enterprise.

Chapter 2 Information in biology

2.1 Introduction and outline

The received view in biology includes the idea that DNA carries information about the phenotype (Alberts et al., 2008; Futuyma, 1998; A. J. F. Griffiths, 2008; Maynard Smith, 2000; Mayr, 2004; Ridley, 2004). To take a representative example, a leading textbook on cell biology states that cells "use raw materials to create a network of catalysed reactions… according to an elaborate set of instructions encoded in the hereditary information" (Alberts et al., 2008, p. 401). When commenting on "[t]he colloquial use of informational terms" in molecular biology, the eminent evolutionary biologist John Maynard Smith wrote he was "not aware of any confusion arising because their meanings are not understood. In fact, the similarities between their meanings when referring to human communication and genetics are surprisingly close" (Maynard Smith, 2000, p. 178). Philosophers of biology, on the other hand, seem less optimistic, and have repeatedly argued that such information talk, though pervasive, is highly problematic because it encourages genetic determinism, essentialism, and the nature-nurture disjunction (Godfrey-Smith, 2007; P. E. Griffiths, 2001; Oyama, 1985/2000; Sarkar, 1996; Sterelny & Griffiths, 1999). This chapter explores this tension. More precisely, my goal is to provide critical answers to two related questions:

- *Are biologists justified in using information talk to describe and explain molecular and developmental processes?*
- *Are biologists justified in attributing a special informational role to genes in molecular and developmental processes?*

Defining gene is not straightforward (Keller, 2000; Neumann-Held & Rehmann-Sutter, 2006) but in this piece I assume an inclusive textbook definition of gene as "the

segment of DNA sequence corresponding to a single protein or set of alternative protein variants (or to a single catalytic or structural RNA molecule for those genes that produce RNA but not protein)" (Alberts et al., 2008, p. 7).

As for 'information,' different notions coexist in the biological literature and their uses can be mutually inconsistent. The term often goes undefined and unindexed, and this does not help clarify its meanings across texts. Thus, to address these questions satisfactorily we must clarify what one might mean by 'information' in general and in the context of biology. I will distinguish between the *communication-theoretic* concept of the amount of information in a signal as the reduction of uncertainty associated with it (Shannon, 1948a); the *situation-theoretic* concept of the specific information carried by a signal as nomic dependence (Dretske, 1981) or constraint (Israel & Perry, 1990) between situations; the *teleosemantic* concept of information as biological function resulting from natural selection (Maynard Smith, 2000; Sterelny, Smith, & Dickison, 1996), and the concept of information as *a difference that makes a difference* for the system under study (Bateson, 1972; Oyama, 1985/2000).

In my discussion, I shall mention the roles genes play in molecular and developmental processes. In doing so, I will not dispute the conventional body of scientific knowledge and empirical observations in molecular biology, cell biology, and genetics, such as the model for the molecular structure of the DNA initially suggested by Watson and Crick (1953a, 1953b), the processes involving DNA which result in the production of mRNA (transcription) and proteins (translation), the reliable correlations that are observed between the sequence of nucleotides in the DNA and the sequence of amino acids in the corresponding polypeptide (the 'genetic code'), or the measurable correlations between alleles (or in fact any feature of the developmental complex) and other phenotypic features. These will be taken as well-constructed scientific facts (Alberts et al., 2008; A. J. F. Griffiths, 2008). What will be open to critical evaluation are the questions of whether there are justifiable uses of informational concepts in these contexts, and whether the role genes play in development justify the common information talk.

Views about how organisms come to be are relevant not only for the molecular biologist working at the bench, but they also inform how psychologists study cognitive processes. On top of clarifying the use of information among biologists, another goal of

the chapter is to make explicit some of the motivations for rejecting the received computational-representational model of the mind (Fodor, 1975). This model is based on the premise that cognition is about computation with representations and that the nervous system is fundamentally an input-output system that processes 'information' – *which* sense of information is a question I will explore in the following chapters. Thus, the present discussion about 'information' in development is an essential step to our subsequent discussions about cognition, social learning and the relation between biology and culture.

The chapter is organized as follows. Because of its central importance to our purpose, I will first distinguish between the concept of information developed in the mathematical theory of communication (Shannon, 1948a, 1948b) and the related, semantic concept of the *specific* informational content carried by a signal (Dretske, 1981). I will then discuss these notions in relation to the roles played by genes and environment in developmental processes. As we go along, it will become necessary to introduce the concept of *intentional* information to address critical points raised by Sterelny and Griffiths (1999) and Maynard Smith (2000). I will suggest that the idea that genes play a special informational role in development is motivated by an underlying, implicit or explicit, way of thinking about how form arises in nature that is 'hylomorphic' (Ingold, 2013) and 'preformationist' (Oyama, 1985/2000, 2000b, 2015). These ideas, although present in influential textbooks and part of the received view, do not find support in empirical facts and biological theory (Sarkar, 1996) and I will indicate an alternative way of thinking that motivates this thesis. At this point I will present yet another sense of information as a difference that makes a difference (Bateson, 1972; Oyama, 1985/2000). For clarity, I will always make explicit (usually in a note within brackets) *which* concept of information is being referred to when I use the term. I think this is good practice that might help avoid misunderstanding and equivocation. I will also occasionally use 'information' with single quotation marks when no specific sense is evoked.

2.2 What is information? An initial distinction

2.2.1 Information as reduction of uncertainty in the mathematical theory of communication

The initial development of the mathematical concept of information is associated with an interest in solving the technical problems faced by engineers working with the transmission of signals through telegraph and later telephone lines (Markowsky, 2013). It was in this context that Claude Shannon published his mathematical theory of communication in two articles in the technical journal of the Bell Telephone Laboratories where he worked (Shannon, 1948a, 1948b). Shannon's work provided engineers with a systematic way to analyse specific communication systems and, for example, improve performance in signal transmission in existing telephone cables. He developed a simple and fairly abstract model of a communication system which includes signal (or message, state, and so on), noise, source, transmitter (or encoder), channel, receiver (or decoder), and destination. Because of its abstract formulation, we can think of the *source* as a person, an animal, a computer, a thing or any process which selects (intentionally or metaphorically) a *signal* from a set of possible alternatives. The *transmitter* or encoder transforms or encodes the message into a form suitable for transmission through the channel. The *channel* is the medium or the set of physical-chemical processes that transmits the signal to the receiver. The transmission of signals through the channel can be more or less affected by sources of *noise*. The *receiver* or decoder transforms or decodes the signal into a message that can be utilised by the destination. The *destination* can also be a person, an animal, a computer, a thing or process which interprets or otherwise uses the message.

Before we consider the *flow of information* in communication systems, we must clarify how information is generated in the first place. According to Shannon (1948a), information is generated when a message is chosen from a set of alternatives. The process of selecting a state from a set of possibilities is associated with some *reduction of uncertainty* – say reducing from a set of 100 possible messages to the one that is actually chosen, encoded and transmitted. Shannon proposed a way to quantify this process and called it a measure of the *amount of information* associated with the signal generated at the source. This quantification has nothing to do with either the *meaning* or the *specific information* a message may convey to someone or something able to interpret it.

One way to reduce a set of alternatives to one is by asking a series of yes/no questions. The answer to each such question distinguishes the initial set into ever smaller groups which will eventually include only the one chosen message. Each answer can be represented by a binary digit or bit (e.g. 1 for yes and 0 for no, or the alternative states in a transistor) and thus each message in the set can be represented by a unique combination of the bits which represents the answers to such questions (e.g. 1010001). Therefore, we can use the number of the binary decisions required to identify the message at the source, s_i, as a measure of the amount of information (*sensu* Shannon entropy or reduction of uncertainty) associated with it, I. Following previous suggestions, Shannon defined this quantity as the negative logarithm (of base 2 for binary decisions) of the probability that a message is chosen by the source, $p(s_i)$, and is formalized as follows:

$$I(s_i) = -\log_2 p(s_i)$$

Engineers are less concerned with particular messages and more with the average amount of information generated by the source and by the receiver, and especially with how they are coupled. The average amount of information produced by the source, $I(s)$, depends on the probability that each possible message is actually selected, $p(s_i)$, and the amount of information each contains. This is given by the formula:

$$I(s) = -\sum p(s_i)\ \log_2 p(s_i)$$

The same formula can be used to compute the average amount of information generated in the receiver. Then one can calculate, for example, how much information in the receiver correlates with information in the source and how much does not. In sum, according to the mathematical theory of communication, the *amount of information* carried by a specific message is a measure of the reduction of uncertainty associated with its being chosen, which is relative to the alternatives being distinguished.

One important feature of Shannon's treatment is that he focuses mainly on the problems related to the encoding of signals (transforming the chosen messages into a form suitable for the channel) and their transmission. He explicitly avoids semantic aspects of communication such as how signals can have meaning and be about something else: "These semantic aspects of communication are irrelevant to the engineering problem. The

significant aspect is that the actual message is one selected from a set of possible messages" (Shannon, 1948a, p. 379).

When you (the source) are talking to a friend (the destination) on the phone the sounds you make correlate with the electric signals generated in your device (the transmitter), which in turn correlate with the electric signals and thus the sounds generated in your friend's device (the receiver and decoder) as constrained by the channel conditions linking them. As a result, your friend (the destination) can then try to make sense of what she hears. In this case we can say that some *amount of information* (*sensu* Shannon entropy or reduction of uncertainty) is transmitted between your mobiles. But this notion of information does not refer to the interpretation of the meaning of the message, which depends on sense-making abilities supported by linguistic conventions and the context in which communication occurs. Neither does it refer to what you intended to convey, or what your friend now thinks you meant. Anyone who has tried to communicate their ideas, or who tries to understand a message spoken or written in an unfamiliar language or about an unfamiliar topic, will have experienced that the meaning intended by the speaker and the meaning constructed by the listener can be frustratingly different. Shannon's concept of information does not refer to what you might learn, from the signal or message, about other events or processes, regardless of the communicator's intentions. To address these aspects of information, we must distinguish between the communication-theoretic concept of the *amount of information* (*sensu* Shannon entropy or reduction of uncertainty) and the semantic concept of the *specific informational content*, or *specific information*, associated with a message. This distinction was advanced by Dretske (1981) and is also developed in situation theory (Barwise & Perry, 1981; Devlin, 2006; Israel & Perry, 1990).

2.2.2 'Information about' as constraint between situations in situation theory

Meaning and intentionality have been contentious topics in linguistics and the philosophy of mind, but they are central to the ordinary use of the term information (P. Jacob, 2014). People (including philosophers) say, for example, that perceptions, thoughts, and feelings, and also the sentences spoken to communicate them, are *about* features external to the 'mind'. Intentionality or reference to an object has famously been considered to be

a “characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it” (Brentano, 1995, pp. 88-89, originally published 1874). More recently, philosophers have attempted to naturalize the notion (e.g., Barwise & Perry, 1981; Dennett, 1987; Dretske, 1981; Millikan, 1984).

Intentionality and (biological) meaning are also central to the way the average biologist ordinarily uses the term, such as in reference to the metaphor of a genetic code, the correlations between segments of DNA and phenotypic features, and the correlations between features in the environment and the behavioural or cognitive states of the perceiving-acting organism (Maynard Smith, 2000). In this section I will present some aspects of the semantic theory of information developed by Dretske (1981) and the situation theory as initially developed by Barwise and Perry (1981), both of which build on Shannon’s treatment. These works provide a clear path from the idea that any situation can generate some *amount of information* (*sensu* Shannon entropy or reduction of uncertainty) to the idea that signals or situations carry *specific information about* other signals or situations.

Despite the original context of interest, namely the telephone and other physical systems, the mathematical notion of information can be conceived in far more general terms. “Any situation”, as Fred Dretske explains, “may be taken, in isolation, as a generator of information. Once something happens, we can take what did happen as a reduction of what could have happened to what did happen and obtain an appropriate measure of the amount of information associated with the result” (Dretske, 1981, p. 14).

In addition, the concept of communication channel can also be conceived in more general terms. If the state of affairs or situation *B* depends on the state of affairs or situation *A* due to (lawful) regularities, then we can say that situation *B* carries *specific information about* situation A, and vice versa (Dretske, 1981).

Another way to express this semantic notion is to say that the specific information, or the informational content, of a signal or message is “the *what-it-is-we-can-learn* from that signal or message (in contrast to the *how-much-we-can-learn*)” (Dretske, 1981, p. 47). This is how Dretske (1981, p. 45) defines the semantic concept of informational content:

> A state of affairs contains information about X to just that extent to which a suitably placed observer could learn something about X by consulting it. This, I suggest, is the very same sense in which we speak of books, newspapers, and authorities as containing, or having information about a particular topic.

For example, the volume of a liquid such as ethanol covaries with its surrounding temperature, and this regularity allows us to use ethanol to build thermometers: the values reading in the thermometer carries specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) about the air temperature and vice versa. Thus, the fact that the length of the ethanol column in the thermometer attached to my window is such and such carries the specific information that the air temperature outside is such and such. What that means is that I, the observer, can read the values on the thermometer and learn about how many layers I should wear before leaving the house. Crucially, relative to the constraint, the air temperature carries specific information about the ethanol column as well, by the same token.: If I go for a walk and experience that it is snowing, I can learn something about the value currently reading in the thermometer back home (e.g., that it is probably at or below zero degrees Celsius).

Consider the following question, which was implicitly hinted at above. Is specific information something objective – in the sense that it exists in the world independent from any observer or cognitive agent detecting it, or is specific information always information relative to some observer? The answer is relevant for our discussion because biologists talk about information being *in* the DNA with an objective tone to it (Maynard Smith, 2000). Let me refer back to Shannon (1948a). In his account, information (*sensu* reduction of uncertainty) is *generated* when one possibility is selected among alternative possibilities, and in this abstract formulation it does not depend on any observer. As for the semantic notion developed above, there is an unresolved tension around this point. On the one hand Dretske (1981, p. 57) considers that

> Information, as defined above, is an objective commodity, the sort of thing that can be delivered to, processed by, and transmitted from instruments, gauges, computers, and neurons. It is something that can be in the optic array, on the printed page, carried by a temporal configuration of electrical pulses, and stored on a magnetic disk, and it exists there whether or not anyone appreciates this fact or knows how to extract it. It is

> something that was in this world before we got here. It was, I submit, the raw material out of which minds were manufactured.

On the other hand, he also states that "how much information a signal contains, and hence what [specific] information it carries, depends on what the potential receiver already knows about the various possibilities that exist at the source" (Dretske, 1981, p. 79). What should count, one might ask, as the relevant possibilities being discriminated? "When a possibility becomes a relevant possibility is an issue that is, in part at least, responsive to the interests, purposes, and, yes, values of those with a stake in the communication process." (Dretske, 1981, p. 132). Thus, on the one hand, specific information is "an objective commodity", and on the other hand it is a relative concept that depends on the knowledge state of the information user. Dretske further says that specific information "has this social, relative, pragmatic dimension to it" (Dretske, 1981, p. 132), and the same applies to knowledge, which he defines as "information-produced belief" (Dretske, 1981, p. 92).

Dretske acknowledges the tension but does not see it as problematic for his purpose. "The relative character of information", he says, "does not make it 'less objective' but it requires that the receiver to which one refers is made clear" (Dretske, 1981, p. 79). One way to escape this tension is by resorting to what he calls a harmless fiction: "we have indulged in the harmless fiction that the number of possibilities existing at the source (and their respective probabilities) was fixed independently of what anyone happened to know" (Dretske, 1981, p. 57). His specific suggestion is that information *is* a relative concept but it is harmless to speak *as if* it were an absolute concept in such cases where "[e]very relevant party knows the same thing about the possibilities existing at the source" (Dretske, 1981, p. 80). "Only when there is a shift of reference systems", he argues, "does the need arise to make explicit the relative nature of the quantity under consideration" (Dretske, 1981, p. 80). This may not be problematic for his purpose, but we must remember that his purpose is to provide a semantic theory of information that is useful for the domains of cognition and language use. He is not trying to account for the role genes play in development or protein synthesis but things like belief, knowledge, concept, and meaning; as well as sensory, cognitive, and learning processes. These processes imply cognitive systems detecting and making use of the 'information' available in the environment via action and perception (see next chapter).

Situation theorists write in a similar spirit. “What underlies the phenomenon of information”, say Israel and Perry (1990, p. 3), “is the fact that reality is lawlike; that what is going in one part of reality is related to what is going on in some other part of reality, by laws, nomic regularities, or as we shall say, *constraints*”. Because the world is “knitted together by constraints… situations carry information. The fact that there is a situation of one type, carries the [specific] information that there are situations of the types that one involves” (Israel & Perry, 1990, p. 3). This passage refers to constraints as lawful regularities, as did Dretske (1981) but other authors suggest a more inclusive notion. “Constraints are linkages between situation types. They may be natural laws, conventions, logical (i.e., analytic) rules, linguistic rules, empirical, law-like correspondences, etc.”(Devlin, 2006, p. 608).

Note that this introduces the relation of intentionality or ‘aboutness’ into the picture, since the informational content conveyed by a situation is *about* some other situation. However, this relation is not between ‘mental phenomena’ and ‘physical phenomena’ as in Brentano (1995), but between situations (parts of the world). Thus, intentionality need not imply a distinction between the body and the mind.

So far, I have started to clear the conceptual ground and distinguished between two concepts: the amount of information (*sensu* Shannon entropy or reduction of uncertainty) associated with a signal and the specific information it conveys (*sensu* constraint between situations or what-it-is-we-can-learn-from-it). I will now finally move to our scientific context of interest and examine information talk in biology.

2.3 Information talk in biology

Philosopher of biology Sahotra Sarkar (1996) suggests that the term information was introduced in the molecular biology literature by James Watson and Francis Crick in their seminal discussion of the implications of their model for the structure of the DNA, where they (Watson & Crick, 1953a, p. 965) say:

> The phosphate sugar backbone of our model is completely regular, but any sequence of the pairs of bases can fit into the structure. It follows that in a long molecule many different permutations are possible, and it therefore seems likely that the precise sequence of the bases is the code which carries the genetical information.

The idea that genes carry information *about* proteins or the phenotype has since become widespread. It is indeed possible to talk about the process of protein synthesis, or even the entire development or ontogeny of organisms, using a version of Shannon's model of communication (for a review and explicit defence of this talk in biology, see Maynard Smith, 2000). The DNA (or mRNA, it does not matter for the analogy) is the transmitter, the protein or phenotype is the receiver, and all non-genetic components of the developmental matrix, along with the constructive processes they take part in, constitute the constraints or channel conditions – which usually get lumped under the vague rubric 'the environment'.

Suppose I follow a common experimental design and raise organisms with different alleles in very similar conditions; and suppose I find that different genotypes correlate with some aspect of the phenotype. *If I ignore the complexities of actual developmental processes*, then I might say that some amount of information (*sensu* Shannon entropy or reduction of uncertainty) was transmitted *from genotype to phenotype* by means of development. I might also say, following Dretske (1981) and Israel and Perry (1990), that the genotype carries specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) about the phenotype. I assume this analogy is not necessarily controversial; however, three points are in order.

The first point is that, whenever situation types are connected by constraints, specific information flows between situation tokens in both directions. Therefore, we might say that, *for an observer*, the genotype carries specific information about the phenotype and we might also say that the phenotype carries information about the genotype. As Godfrey-Smith tells us, "the information in the genes is something that only we, the observers and describers of genetic systems, use. This information is not part of any explanation of the biological role that the genes play within organisms. Here, genes are being used by us just as we might use tree rings [to infer the age of the tree]" (Godfrey-Smith, 2004, p. 278).

When the observer is explicitly mentioned, as above, this symmetry in informational relations does not seem particularly problematic: depending on the available prior knowledge about the system under study, a geneticist can learn about the probable phenotype of an organism by consulting its genotype; alternatively, she can

learn about its probable genotype by consulting the phenotype. But in the received view, 'information' is used in a sense that is not consistent with "that commodity capable of yielding knowledge" to an observer (Dretske, 1981, p. 44). Rather, it is conceived of as being in the DNA in some objective or absolute sense, and as encoding *instructions* to create and maintain metabolism. I conclude that this use, so common among biologists, is different from the notion of specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) I have presented above.

The second point is that the distinction between source and channel conditions in a system of causal relations is a matter of convention. Suppose I use a spring scale to measure the mass of unknown materials. Given the instrument's design and an appropriate gravitational field, the amount of mass in the material correlates with the state of the spring and, thus, with the position of the pointer. On the one hand, if the focus of attention is on the material being tested, I might say that the instrument's design and the gravitational field compose the channel conditions and that the pointer (destination) carries specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) about the material (source). On the other hand, during calibration the focus of attention is on the inner functioning of the instrument itself rather than the materials being measured. To calibrate the scale, one must measure materials with *known* values of mass. This allows us to adjust the parameters of the instrument as required. Note, however, that now the pointer (destination) carries specific information about the state of the instrument (now the source) and the materials being measured fall into the background as part of the constraints or channel conditions.

Turning to biology, as well as varying the genotype while holding the external micro-environmental conditions similar, one could also hold the genotype constant and vary the environment, as geneticists routinely do (using clones) or what studies of identical twins to some extent set out to do through statistical procedures. Now in this design different values of the environmental variable may correlate with different phenotypes and I should be equally justified in saying that some amount of information (*sensu* Shannon entropy or reduction of uncertainty) was transmitted from the environment to the phenotype through the channel of development. I could also say that

the environment carries or conveys specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) about the phenotype.

This other symmetry in informational relations has been pointed out by several philosophers of biology in previous discussions on the topic (Godfrey-Smith, 2007; P. E. Griffiths, 2001; Oyama, 1985/2000, 2000a; Sarkar, 1996; Sterelny & Griffiths, 1999). For example, suppose I incubate eggs with clones of the American alligator (*Alligator mississippiensis*) at different temperatures. Depending on the specific temperature, the eggs will develop and become either females or males (Gilbert, 2000). In this case, for an observer with the appropriate background knowledge, the temperature of incubation carries specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) about the sex that one can expect the embryos to develop. If you incubate eggs at 30^{o}C you can predict that all animals will probably become female, and if you know an animal's sex you can infer the approximate range in temperature it experienced during the critical period.

Thus, any developmental resource should count as a legitimate source of specific information about development, and not just the genome. But I think no biologist would conclude, from the premise that the environmental variable under consideration carries specific information about the phenotype, that environmental features encode "an elaborate set of instructions" (Alberts et al., 2008, p. 401) for development. Yet, the notion that DNA does contain instructions is unlikely to raise an eyebrow. Some other premise is in place. I will return to this point later; first I must turn to another strategy to conceptualize information in biology.

2.3.1 The teleosemantic concept of intentional information as biological function

In a comment about Dretske's (1981) notion of specific information, Maynard Smith says it "has the virtue of clarity, but… [it] would rule out the current usage of the concept in biology" (Maynard Smith, 2000, p. 189). The reason for this, he argues, is that Dretske's definition can be legitimately applied to any component that correlates with developmental outcome (as I have pointed out above). "Colloquially, this is fine;" he concedes, "a child's environment does indeed predict its future. But biologists draw a distinction between two types of causal chain, genetic and environmental, or 'nature' and

‘nurture’, for a number of reasons” (Maynard Smith, 2000, p. 189). I will present and discuss his reasons further below (in section 2.3.3). For now, let us follow closely his attempt to justify “the current usage” of the term.

The starting point of his argument is to distinguish between what he calls two different contexts in which the notion of information is used in colloquial speech. The first context he mentions is related to Dretske’s (1981) semantic concept of specific information, or informational content, as nomic dependence: “for example, we may say that the form of a cloud provides information about whether it will rain”. The other context involves intentionality. “In biology,” he says, “the statement that A carries information about B implies that A has the form it does because it carries that information” (Maynard Smith, 2000, p. 189). To clarify the difference between the two contexts he says that a “DNA molecule has a particular sequence because it specifies a particular protein, but a cloud is not black because it predicts rain. This element of intentionality comes from natural selection” (Maynard Smith, 2000, p. 189). His thesis then is that “[t]he notion of information as it is used in biology… implies intentionality. It is for this reason that we speak of genes carrying information during development, and of environmental fluctuations not doing so” (Maynard Smith, 2000, p. 193).

His reference to intentionality deserves close attention. The only work referenced in this passage comes not from biology but from philosophy of mind. The cited work is *The intentional stance*, by philosopher Daniel Dennett (1987), in which he distinguishes between three levels of abstraction in the study of the mind – the physical stance, the design stance, and the intentional stance – while famously advocating for the intentional stance as a strategy to understand human behaviour. In this context intentionality is a property of mental states such as believing and thinking. The domain of inquiry discussed in this work is quite remote from the molecular processes in which genes participate. To justify the attribution of intentionality to the genome, Maynard Smith uses an analogy with genetic algorithm, an optimization strategy used in software development inspired by the very notion of evolution by natural selection. In the implementation of genetic algorithm different codes ‘compete’ against each other in their ability to solve a predefined computational problem. They undergo an iterative procedure involving ‘selection’ of the best scripts and ‘mutation’ of these scripts over ‘generations’ until some

code meets the satisfaction criteria. Just as "there is information in the bit string, which has been programmed by selection [in the genetic algorithm]", he suggests that the same is true for the information in the genes, arguing that "[t]his element of intentionality [present in the genes] comes from natural selection" (Maynard Smith, 2000, p. 190). Thus, he concludes that

> the analogy… [between genetic algorithm and evolution by natural selection] justifies biologists in saying that DNA contains information that has been programmed by natural selection; that this information codes for the amino acid sequence of proteins; that, in a much less well understood sense, the DNA and proteins carry instructions, or a program, for the development of the organism; that natural selection of organisms alters the information in the genome; and finally, that genomic information is 'meaningful' in that it generates an organism able to survive in the environment in which selection has acted.

Let us follow him a bit further. He says that a "very similar conclusion about the concept of information in biology has been reached by Sterelny and Griffiths (1999)". The passage he then quotes from these authors seems to support this: "In particular, they [Sterelny and Griffiths (1999)] write, 'Intentional information seems like a better candidate for the sense in which genes carry developmental information and nothing else does'" (Maynard Smith, 2000, p. 193). As P. E. Griffiths (2001) points out, the strategy to define intentional information in terms of biological function derived from a history of selection relates to the teleosemantic strategy to naturalize 'meaning' (Millikan, 1984; Papineau, 1987). One well-developed version of the teleosemantic strategy in biology argues that it is legitimate to talk about a genome representing or carrying 'information' about the developed phenotype because "the genes have the biofunction of guiding phenotypic development" (Sterelny et al., 1996, p. 388).

Interestingly Maynard Smith fails to acknowledge a critical point that Sterelny and Griffiths (1999) developed in the paragraphs just following the quoted passage, regarding attributing intentional information to genes alone: "many other means through which parents influence their offspring have selection histories too" (Sterelny & Griffiths, 1999, p. 105). The authors argue that, if one wants to use the teleosemantic concept of intentional information, it should apply to "all adapted developmental resources" whose present form results from a history of selection (Sterelny, 2000, p. 197). Some authors

refer to these resources, which may be reliably available or passed on from parents to offspring, as additional, non-genetic inheritance systems. Suggestions are far reaching and include cytoplasmic factors, obligatory endosymbionts, environmental features, social interactions and learning opportunities (Jablonka, 1994; Jablonka & Lamb, 2005; Odling-Smee, Laland, & Feldman, 1996; Odling-Smee et al., 2003; Richerson & Boyd, 2005).

According to Godfrey-Smith (2007), one problem with the teleosemantic account is that it is not clear why semantic properties should be equated with biological function: "having a function in the rich historical sense is not generally sufficient for having semantic properties. Legs are for walking, but they do not represent walking. Enzymes are for catalysing reactions, but they do not instruct this activity" (Godfrey-Smith, 2007, p. 108).

Another problem with Maynard Smith's argument, as I see it, is the following. Being directed more to justifying current usage rather than to provide a critical discussion, he takes for granted the received idea that DNA contains information in the sense of instruction or specification of form: "the genome is not a description of the adult form, but a set of instructions on how to make it: it is a recipe, not a blueprint" (Maynard Smith, 2000, p. 187). The idea "that the genome contains enough information to specify the form of the adult", he suggests, "is a reasonable assumption, because it is hard to see where else the information is coming from" (Maynard Smith, 2000, p. 186). With this as a premise, he asks the, as he puts it, "hard question. If there is 'information' in DNA, copied to RNA, how did it get there?" (Maynard Smith, 2000, p. 180). His answer, as I pointed out, is that natural selection put it there. But, as I argued above, whatever *clear* definition of information one chooses, there is no empirically or theoretically justifiable way to find it only in the DNA but not in other components of the developmental matrix that also meet the defining criteria.

What is puzzling is that the philosophers' conclusion seems to be unwelcome, unpleasant, and uncomfortable to a lot of reasonable scientists such as Maynard Smith. It would seem that this is not the conclusion biologists would want philosophers to arrive at. But this is indeed the conclusion we have in front of us. Why is it that we find it so difficult to let go off the unquestioned idea that there *must* be information in the genes in

the sense of instructions on how to make organisms? Why is it so "hard to see" (Maynard Smith, 2000, p. 186) things from a different perspective? I suggest two reasons. The first reason is the well-constructed biological fact known as the genetic code. The second reason is a specific way of thinking about biological phenomena. I will examine each in turn.

2.3.2 The genetic code

The expression 'genetic code' refers to the reliable correlation that usually holds between the sequences of triplets of bases in the DNA – or, more accurately, in the messenger RNA (mRNA) – and the sequences of amino acids in the proteins they take part in producing. Proteins are synthesized in the cells within a complex network of chemical processes that involve a great many number of enzymes, ribonucleotides, ATP, ribosomes, DNA, transport RNA (tRNA), messenger RNA, amino acids, etc, all in a solution with the appropriate pH, pressure, and temperature. Given the typical chemical composition of a cell, specific triplets of bases in the DNA correlate with specific triplets of bases (codons) in the mRNA and with specific amino acids in the protein, but if any of these processes fail to obtain or is altered, no protein will be produced or the end product might be altered (Alberts et al., 2008).

The same set of predictable relations is observed in almost all living beings studied so far. For example, the triplet TAT in the transcribed portion of the DNA correlates with the codon AUA in the mRNA and this in turn correlates with an Isoleucine residue in the corresponding protein. These reliable associations can be represented in a table of correlations which is called the genetic code. The code is thus a *condensed narrative* about these molecular processes, about their reliability and robustness within organisms, about their similarity across taxa, and about the scientific practices that justify referring to them as a well-constructed fact in biology.

In Dretske's information-theoretic terms, for a suitably positioned observer with the relevant background knowledge, the fact that a sequence of bases in a DNA is such and such carries the specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) that the expected sequence of amino acids is such and such. This does not imply or support the idea that the DNA carries instructions or programs that

control development, but only that DNA sequences covary with phenotypic outcomes and thus may serve as a more or less reliable index for the developmental processes that construct them. Reliable correlations such as these exist because the world is "knitted together by constraints" (Israel & Perry, 1990, p. 3) and might extend far beyond proteins to any aspect of the phenotype being tracked.

The robust correlation between sequence of bases in the DNA and sequence of amino acids in the corresponding protein is a good candidate for the attribution of some *special* roles to genes. For example, Godfrey-Smith suggests that "the crucial features of gene expression mechanisms… motivate the introduction of a symbolic or semantic mode of description" (Godfrey-Smith, 2007, p. 110). Furthermore he argues that the reliable correlation, which is commonly referred to as genes *coding for* the primary structure of proteins, "is one kind of informational or semantic property that genes and only genes have" (Godfrey-Smith, 2007, p. 110). It is not entirely clear what sense of information is being referred to in this passage but in other passages it is referred to as a "richer sense… sometimes called 'semantic' or 'intentional' information (Godfrey-Smith, 2007, p. 106) and associated with the teleosemantic strategy I mentioned above (Sterelny et al., 1996). In any case, he argues that the fact that genes reliably correlate with (or, in the common parlance, 'code for') the sequence of amino acids *does* justify attributing intentional information to genes but *does not* justify the promiscuous use of the term in other contexts: "this relation [of coding or specificity] 'reaches' only as far as the amino acid sequence. It does not vindicate the idea that genes code for whole-organism phenotypes, let alone provide a basis for the wholesale use of informational or semantic language in biology" (Godfrey-Smith, 2007, p. 110). Sarkar (2004) makes a similar point.

I suggest that, though the table of correlations or genetic code is a well-constructed fact, the attribution of intentionality and meaning to DNA is not straightforward. This is especially clear in relation to eukaryotes. I disagree with Godfrey-Smith's suggestion that the complicated picture of protein synthesis in eukaryotes can be ignored as "mere details" (Godfrey-Smith, 2007, p. 110) that do not affect the attribution of intentional information to genes alone, and in support I cite four, equally well-constructed, textbook facts: (1) the 'universal' genetic code is not really universal, (2) the same DNA transcript can undergo alternative splicing depending on the local molecular context, (3) the

function of the polypeptides depend on more than their linear structure, and (4) properties of complex systems such as cells and multicellular organisms cannot be reduced to the properties of lower-level components (see below). All these facts blur the neat picture of DNA correlating with (or 'coding for') proteins as a special informational relation. I will expand on each in turn.

Firstly, exceptions to the so-called universal genetic code are known in many taxa, especially in their mitochondria. For example, in vertebrates the triplet AUA in a mitochondrial mRNA that is translated inside the organelle (rather than in the cytoplasm) will correlate with Methionine instead of Isoleucine (Osawa, Jukes, Watanabe, & Muto, 1992). The reason is because the mitochondrial tRNA that attaches to Methionine has a slightly different anticodon structure (Takemoto et al., 2009). In other words, imagine I conduct an experiment in which I put copies of the same DNA sequence to 'code for' proteins in two conditions: in the cytoplasm and in the mitochondria. What this design effectively does is keep the sequence of DNA bases constant while changing the chemical context. Once you know the location of transcription and translation, you can look at the appropriate table (either the 'universal' cytoplasmic code or the vertebrate mitochondrial code) and predict the linear structure of the protein that will be produced in each case. In this scenario, it is the location rather than the DNA sequence that covaries with the protein and therefore carries specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) about its structure. The sequence of bases in the DNA is an invariant contextual condition that is a part of the 'channel condition'. If DNA carried semantic information in some absolute sense, its meaning would not change depending on the surrounding molecular context.

Secondly, consider alternative splicing. In eukaryotes (a major clade that includes all organisms with nucleated cells such as plants, animals, fungi, and protists), when DNA is 'transcribed', the initial RNA is further transformed ('edited') by chemical processes in which some parts (called 'introns') are removed and the others ('exons') are joined and become the mRNA. Which parts get removed and which are kept is context-dependent. Thus, within the same cell, the same stretch of DNA may correlate with several alternative mRNAs and proteins depending on what is going on in the rest of the cell at that point in

its ontogenetic history. As above, if DNA carried 'information' in some objective sense, its content should not change depending on the surrounding molecular context.

Thirdly, the linear sequence of bases is a fundamental aspect of a protein but ultimately its function depends on its three-dimensional shape, and sometimes subunits aggregate to form the functional protein complex. These folding processes depend critically on the chemical context in which polypeptides emerge and may involve further modifications by other enzymes (e.g. glycosylation or phosphorylation). These folding processes are not specified as instructions in the sequence of bases and yet they are fundamental for the appearance of the final, functional form of the corresponding protein. Therefore, they have a legitimate role to play in causal accounts.

Fourthly, the fact that organisms are internally heterogeneous brings about the question of how to conceive the relations among a collection of molecules and the whole organism they compose. Zoom in on a digital picture of a lion and you see a collection of static, discrete pixels that together compose the complete image. Zoom in on the lion itself and what you see (or what I imagine you would see according to current scientific models) is not a static spatial composition of molecules but a buzzing network of chemical transformations which *correspond to* the different, ever-changing structures of the lion as described at different scales: cells, tissues, organs, and the lion's whole body. Now move your attention to the furry skin of the animal and what you see is not a solid surface but a dynamically stable boundary that is continually being constituted as molecules flow in and out of cells, as the animal moves about perceiving and acting in its environment.

How one conceives the relations among scales and 'levels of analysis' is far from an easy question (DiFrisco, 2017). Yet it also constrains the kinds of narratives one is likely to produce about genotype- phenotype relations. If we start from the notion that all higher level features of the organism are reducible to the lower level properties of its molecules (the received reductionist position), than we are arguably more likely to produce narratives in which genes play a central role in development, heredity, and evolution (such as in Alberts et al., 2008; A. J. F. Griffiths, 2008; Maynard Smith, 2000; Mayr, 2004; Ridley, 2004); and in which the phenotype is seen as "the physical manifestation of a genotype (gene combination)" (Futuyma, 1998, p. 37). On the other hand, if we conceive of organisms as complex systems, as multiscale processes affecting

each other, then we are arguably more likely to produce narratives in which no single scale and no single class of components have primacy; in which the "gene combination" is not a set of instructions for the phenotype but rather one of its components; in which a shift from lower to higher level processes is conceived, not as a change from causes to effects, but rather as a change in "scale, vocabulary, concepts, and method" (Oyama, 1985/2000, p. 164) and as a matter of finding possible correspondences (Gottlieb, 2007; Kauffman, 1993; Levin, 2006; Maturana & Varela, 1980, 1992; Oyama, 1985/2000; Rose, 1981; Thelen & Smith, 1994).

The specificity in the relations between DNA and amino acids in proteins is what grounds the concept of a 'genetic code' and arguably what motivates the attribution of a special informational role to genes. The general conclusion I want to draw from the textbook facts briefly reviewed above is that this specificity cannot be attributed to the linear sequence of bases in the DNA structure *independently* of the biosynthetic processes they take part in. These processes are made possible by the structure and properties of both the DNA and the rest of the relevant chemical context, and this context is itself a dynamic process rather than a static set of 'raw materials' or 'enabling conditions'.

The 'genetic code' is an epistemic abstraction, a well-constructed fact and a very useful and compact representation of the complex relations involving DNA, mRNA, the produced polypeptide, and so many other molecules. These relations should not be reified – they are not substantively *there* but are continually being generated in the cells. This is far from saying something like 'DNA does not matter' for the structure of the final protein because of course it does. This is also far from saying that 'all components play the same role' because of course they don't. DNA plays a different role than, say, RNA polymerase, in protein synthesis. This is also far from ignoring that sequences of DNA (as well as other components of the developmental matrix) have a rich history of 'selection' that explains their current form.

What this *does* imply is that there are no logical grounds on which to support the statement that DNA sequences carry 'information' in the sense of instructions, specifications, programs, or some other form of nature's intention, that is independent from the very biosynthetic process in which they participate, and also from the epistemic relations established by the questions the experimenter is trying to address. To say that

the role DNA plays in development is that of the commander-in-chief, the locus of causal control, or the source of biological form is not only unnecessary to convey the idea of contingent specificity. It is empirically and theoretically unjustified.

I am in good company here. "There are numerous important differences between what DNA does in development and the roles played by other causal factors", says P. E. Griffiths (2001, p. 410), "but these differences do not map onto a distinction between informational and material causation. The present atmosphere, in which information talk is only applied to genes… misleads people about the forms of explanation in molecular biology". In a similar tone, Godfrey-Smith (2007, p. 112) agrees "that there is something definite about informational description of genes that encourages fallacies about genetic causation." Something is lurking behind such that "the inferential habits and associations that tend to go along with the use of informational or semantic concepts lead us to think of genes as having an additional and subtle kind of extra causal specificity" (Godfrey-Smith, 2007, p. 113).

Following Oyama (1985/2000, 2000b, 2015), I suggest that this asymmetry in how genes, compared to everything else, figure in information talk in biology has to do with a certain view of development, inheritance, and evolution that usually goes unquestioned and mostly even unsaid as part of the received view. This is a way of thinking about biological phenomena which starts from the premise that genes play a special causal role *because* they carry instructions or encode programs. This is a way of thinking in 'nature' and 'nurture' continue to be invoked even as authors indicate the inadequacy of the opposition. I will refer to this special role assigned to the genome as the "source of form", for reasons that should become clear as we go along. I now must pick up a loose thread I had put to the side some time ago, and examine the reasons Maynard Smith gives to why "biologists draw a distinction between two types of causal chain, genetic and environmental, or 'nature' and 'nurture'" (Maynard Smith, 2000, p. 189).

2.3.3 The colloquial use of 'information' as 'instruction' or 'source of form'

I had mentioned above that, in his attempt to justify current information talk in biology, Maynard Smith (2000) lists three reasons why the nature-nurture distinction has become

fundamental in biology. The reasons he gives, which are representative of the received view, are as follows (all quotes below taken from Maynard Smith, 2000, p. 189):

1. "Differences due to nature are likely to be inherited, whereas those due to nurture are not"
2. "evolutionary changes are changes in nature, not nurture"
3. "traits that adapt an organism to its environment are likely to be due to nature"

One way to criticize the reasons listed above is to acknowledge the growing literature on topics which stress the critical relevance of the 'nurture' side in the opposing pair to inheritance, development, and evolution. I will focus on social learning and 'culture' given the context of this thesis. Empirical research, and mathematical and computational models, provide strong evidence that 'cultural traits', i.e. traits acquired through social learning, can 'spread' among individuals in a population and be inherited through countless generations (Fragaszy & Perry, 2003; Hoppitt & Laland, 2013; Laland & Galef, 2009; Richerson & Boyd, 2005; van der Post & Hogeweg, 2009; van Schaik et al., 2003; Whitehead & Rendell, 2015; Whiten et al., 1999). In other words, individuals growing in different social groups are likely to inherit differences due to nurture (cultural information) *as well as* differences due to nature (genetic information). This calls for an expanded notion of inheritance that questions the usually unquestioned premise that only genes play a special role as the carriers of information (*sensu* instruction, specification, or source of form, see below for the latter term). Now genes *and something else* act as carriers of such 'information'. For example, Richerson and Boyd (2005, p. 61) suggest that "the main objects… capable of storing [the] information [that exists in every culture] are human brains and human genes".

'Cultural traits', which are aspects of the phenotype, can change across generations within an evolving population, and this can lead to a process of co-evolution between genes and 'culture' (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981; Feldman & Laland, 1996; Richerson & Boyd, 2005). In other words, evolutionary changes involve changes in nature *and* changes in nurture. This calls for an expanded notion of evolution that questions the usually unquestioned premise that only genes play a special role in evolution since they are the only carriers of the information (*sensu* instruction,

specification, or source of form) that generates the evolutionarily relevant aspects of the organisms.

Furthermore, there is also strong empirical evidence that behavioural patterns that develop within social relations can help organisms cope with, and species adapt to, their environment. Consider, for example, that many singing birds depend on social interaction to learn the local song dialect, which is crucial to finding a conspecific mate (Freeberg, 2000), or that learning new foraging techniques from social partners can lead to the exploitation of otherwise ignored food items in primates (Navarrete, Reader, Street, Whalen, & Laland, 2016; van Schaik et al., 2003; Whiten et al., 1999). In other words, traits that play a role in adjusting the organism to its environment can be due to nature or due to nurture.

Thus, studies of social learning and culture bring into question the primacy given to the 'nature' side of the opposition by stressing the importance of complementary processes that fall on the 'nurture' side. This literature suggests that the form of the adult phenotype, including its behavioural tendencies, 'comes from' information (*sensu* instruction, specification, or source of form) carried by the genes ('nature') and information acquired through social learning ('nurture'). Note that 'nature' and 'nurture', though complementary to each other and interacting in some sense, are still seen as two domains or two potential sources of information (*sensu* instruction, specification, or source of form). Thus, this way to criticize the reasons why the nature-nurture distinction is fundamental does not lead us much far away from the conventional thinking.

The second way to criticize Maynard Smith's defence of the nature-nurture distinction as unproblematic is to question the premise taken for granted in the received view. The premise in question is that genes *do* play a special role in development, inheritance, and evolution, *because* they carry (most if not all) the information (*sensu* instruction, specification, or source of form) that builds the organism. Susan Oyama (1985/2000, 2000b, 2015) has thoroughly argued that the ordinary use of 'information' in biology reveals an underlying preformationist way of thinking. It is preformationist in the sense that it conceives of development as the unfolding of a form that is already there: if not in some homunculus, then encoded in the genome as instruction, specification, blueprint, or program. In contemporary parlance it is 'information' in this instructional

sense that accounts for the organization of mere formless matter into a living organism (such as in Alberts et al., 2008; Maynard Smith, 2000).

Where did the information specifying the form come from? This is a common question in the context of investigating how morphological as well as behavioural form comes to be. I have pointed out above that many find it "hard to see where else the information is coming from" if not from the genome (Maynard Smith, 2000, p. 186) or from the genome *and* the brains of other organisms through 'cultural transmission' (Cavalli-Sforza & Feldman, 1981; Richerson & Boyd, 2005). This ordinary distinction between two complementary sources of information (*sensu* instruction, specification, or source of form) produces and is in turn supported by the nature-nurture complex of oppositions (genes-environment, biology-culture, instinct-learning, and so on). In philosophy, the idea that objects and living beings are a compound of form (*hyle*) and matter (*morphe*) is associated with Aristotle and is called hylomorphism (Ainsworth, 2016). Thus, a version of hylomorphism and of preformationism seem to be present in the received view.

How do we break this established habit of explaining form by pointing to a previous instance of the very form? The alternative I endorse is to conceive of form as residing in the phenotype itself and not in some prior instruction for it, and to explain form in terms of *form-generating* rather than *form-imposing* processes. This alternative involves shifting away from a view of causality and control in terms of linear relations between antecedent and consequent to one in terms of a multiscale network of causal relations among interactants. Consistent with this shift in mindset, I avoid the view of development as the execution of pre-existing instructions encoded in genetic programs, as well as the view of behaviour as the execution of pre-existing instructions encoded in motor programs. In their place, I embrace a relational, processual view of ontogeny as a multiscale, seamless history of transformations which includes changes in metabolism, morphology, physiology, and behaviour as aspects of each other. Finally, to be consistent with this view of ontogeny, social learning will not be understood as a process of 'transmission of information' but rather as a process of development within fields of social relations. I will say a few words about the first point (joint causation and distributed

control) below. The following points – about ontogeny, behaviour, and social learning – will be developed over the next three chapters.

Genes can bring about changes in the organism of which they are a part by perturbing the current dynamics of ongoing, more or less stabilized, molecular processes, not by 'imposing form onto matter'. Therefore, even if triggered by changes in the rate of transcription of some segment of DNA, the control of metabolism and of development cannot be localized 'in' the genome as information (*sensu* instruction, specification, or source of form) but is rather distributed in the whole organism-environment system. Control is constituted in the effective relations that keeps the cell (and the multicellular organism if it is part of one) going, stable in some features while changing in others, as long as these relations maintain its autopoietic organization (Maturana & Varela, 1973, 1980). Control of metabolism and development, therefore, is always systemic, not 'genetic'.

Incidentally, and important to the topic of this thesis, I point out that the same can be said about the control of behaviour. Neuronal activity can bring about changes in the organism by perturbing the ongoing dynamics of neuromuscular networks, not by 'instructing the muscles'. Therefore, even if triggered by changes in the firing rate of some neuronal assembly, the control of behaviour cannot be localized 'in' the brain as information (*sensu* representation, I will return to this in Chapter 4) but is rather distributed in the whole organism-environment system. Control is constituted in the effective relations that keep the animal behaving adequately in its changing environment; it is always systemic, not 'neuronal'. I will return to this in the following chapters.

Adopting this view of control as systemic and distributed, rather than localized, does not preclude one from making distinctions according to questions and scales of interest. The roles played by the linear sequence of nucleotides in the DNA are different from the roles played by the RNA polymerase; the roles played by parental care are different from those played by foraging opportunities, the roles played by neurons are different from those played by muscles, and so on. This understanding allows us to investigate the same phenomena that motivate questions of the kind "Is it nature or nurture?" without committing to unnecessary *a priori* assumptions. Using the term 'information' as part of our descriptions or explanations does not amount to "introducing

some new and special kind of relation or property" but rather to "adopting a particular quantitative framework for describing ordinary correlations or causal connections" (Godfrey-Smith, 2007, p. 106). Or, as Oyama (2009) puts it in a rhetorical question: "Is it by invoking flows of immaterial information that we are most likely to gain understanding of the mechanisms by which these correlations are made, maintained, lost, or reconstructed? Or should we investigate the interactions themselves?"

2.3.4 Information as a difference that makes a difference

Writing in the context of cybernetics, Gregory Bateson defined a unit or 'bit' of information "as a difference which makes a difference. Such a difference, as it travels and undergoes successive transformation in a circuit, is an elementary idea" (Bateson, 1972, p. 321). In a famous passage about the limits of the mind he invites us to consider a man felling a tree with an axe, modifying or correcting each stroke as he goes along. He argues that the control of this emerging behaviour can be thought of as the transmission of information (*sensu* difference that makes a difference) in the circuit composed by "tree-eyes-brain-muscles-axe-stroke-tree", describing this process as "(differences in tree) - (differences in retina) - (differences in brain) - (differences in muscles) - (differences in movement of axe) - (differences in tree), etc." (Bateson, 1972, p. 323). His unorthodox, ground-breaking point is that the mind is not confined to the organism (or brain) as a self or agent performing purposive actions but extends to the environment.

In her seminal book, Oyama (1985/2000) endorses Bateson's definition. In her words: "Information is a difference that makes a difference (Bateson, 1972, p. 315), and what it 'does' or what it means is thus dependent on what is already in place and what alternatives are being distinguished" Oyama (1985/2000, p. 3; note that my quote above from Bateson's definition came from a different edition). The comment that follows the defining statement points to history and context-dependency, two major themes in developmental systems theory. This definition is different from Shannon's (1948a) conception that information (*sensu* Shannon entropy or reduction of uncertainty) is generated when a signal is distinguished from a set of alternatives.

Translating to the present context, let us assume that the world consists of a multi-scale manifold of components and processes – or differences – going on at the same time.

As the developmental system under study undergoes its history of transformations, some components or processes can become relevant for the system; when they do so they become interactants and effectively part of the developmental system. In other words, they *become information for the system*.

I interpret Oyama's endorsement of Bateson's definition as implying that any component or process that is currently affecting or making a difference to the ontogeny of an organism (either sustaining or changing its current trajectory) can be said to constitute information for the organism's developmental system. For example, a gene may be persistently present inside the cells of an organism, but it is not constituted *as information for the system* until and unless and as long as it makes a difference for the developing organism. The difference a gene or any developmental resource makes is not given in its structure independent from the rest of the ecologically embedded organism. This is so because it is the current state of the system (itself a product of its history thus far) that determines what can become relevant for its ongoing dynamics, i.e. what it can affect and what it can be affected by.

Similar to Dretske's notion of specific information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it), information (*sensu* difference that makes a difference) is not substantial but relational – i.e. it refers to the relations among the components that constitute the organism's developmental system. We now have a definition of information (*sensu* difference that makes a difference) that refers to the developmental system itself and not to an external cognitive agent such as a scientist or a partner in a communication event. This is therefore distinct from the concept of specific information as "that commodity capable of yielding knowledge" relative to an observer (Dretske, 1981, p. 44).

Two further aspects that follow from this definition need to be acknowledged because of their relevance in the context of this thesis. First, information (*sensu* difference that makes a difference) can be applied to any causal influence in metabolism and development and does not grant any special in-formative role to genes. Second, information is not some quantity that exists objectively outside the system and is then 'put in' it. Rather, information (*sensu* difference that makes a difference) is *constituted in interaction*.

The idea that information (*sensu* difference that makes a difference) is *constituted in interaction* should not be confused with the idea that information (*sensu* instruction, specification, or representation) is *distributed among factors* (say, some of it encoded in the genes and some encoded in the brain). The former is a relational notion which follows from and invites a systems view of development as a multi-scale constructive process with no central organizer. The latter reifies information and localizes it in things rather than in relations – this would bring back nature and nurture as two distinct domains or causal pathways in development, which I want to avoid.

2.4 Summary and conclusion

This chapter aimed at providing critical answers to the following questions:

- *Are biologists justified in using information talk to describe and explain molecular and developmental processes?*
- *Are biologists justified in attributing a special informational role to genes in molecular and developmental processes?*

To do so I took a long and winding road to clarify different meanings of the contentious term. According to Shannon (1948a), some amount of information (*sensu* Shannon entropy or reduction of uncertainty) is generated when one signal (message, state) is selected from a set of alternatives. This does not tell us anything about the content carried by a specific signal, i.e., what a signal means, is about, or points to.

To deal with aboutness I presented Dretske's (1981) semantic theory of information and, especially, a version of situation theory (Israel & Perry, 1990). This lead me to a way of understanding informational content as a lawful or normative constraint between situations. In this sense, the information carried by some situation is the what-it-is-we-can-learn-from-it about other situations with which it is associated by regularities. With this notion I found a solid foundation for intentional claims such as (1) the fact that the sequence of bases in a segment of DNA is such and such conveys the specific information that the sequence of amino acids in the corresponding polypeptide will be such and such, given the usual constraints; and (2) the fact that an organism carries such and such allele conveys the specific information that it is likely to develop such and such phenotypic feature, given the usual constraints.

Notably, in this semantic sense information 'flows' in both directions and we are equally justified in saying that (1) the fact that some sequence of amino acids in a polypeptide is such and such conveys the specific information that the sequence of bases in the DNA that took part in producing it is such and such, given the usual constraints; and (2) the fact that an organism shows such and such phenotypic feature conveys the specific information that it likely carries such and such allele, given the usual constraints. I also discussed that this concept of information is observer-relative (although we might treat it as an absolute quantity if all persons involved in the communication event share the same background), and that the distinction between source and channel is fundamentally arbitrary. Thus, I concluded that this semantic concept did not justify attributing some special informational role to genes relative to other entities even as I acknowledged that different entities might well play different roles in metabolism and development.

I have briefly examined the teleosemantic strategy (Maynard Smith, 2000; Sterelny, 2000; Sterelny et al., 1996) which equates 'intentional information' with biological function deriving from a history of selection and similarly found no reason to conclude that this concept should apply only to genes and not to other aspects of the phenotype which form and function also result from a history of selection.

As I reviewed these different notions of information I suggested that it is possible to use the term information in biology in consistent ways. However, using the term does not add any new relations to the causal accounts that are of interest, and none of the notions of information justify the idea that genes carry instructions or specifications, as ordinarily seen in biology.

The answer to the first question above is 'Yes, but'. There are several justifiable ways to use information talk in biology. However, to the extent that introducing the polysemic term invites misunderstanding and equivocation, one should be clear about what exactly one means when invoking it. For example, a DNA sequence has a specific, particular structure and thus might be said to generate information (*sensu* Shannon entropy or reduction of uncertainty). On the other hand, sequences of DNA may correlate with sequences of amino acids (or other features of the phenotype) as a result of biochemical constraints, and thus might be said to carry specific information (*sensu*

constraint between situations or what-it-is-we-can-learn-from-it) about these other features for observers with appropriate prior knowledge.

The answer to the second question above is 'No'. I argued that the persistent idea that 'information' in the genes is special (because of its instructive, form-imposing characteristic) follows, not from empirical or theoretical considerations, but rather from a worldview that I characterized as preformationist and hylomorphic, distinguishing form from matter and nature from nurture.

I then questioned the usually unquestioned premise that DNA *must* play a special formative (and in-formative) role as a carrier of instructions and presented my motivation to search for an alternative framework that eschews preformationism and gives appropriate formative weight to all components that take part in form-generating processes. I argued that, while different scientists may focus their attention on different aspects of the multi-scale processes of inheritance, development and evolution, there is no need to commit *a priori* to a framework that might be characterized as hylomorphic, preformationist and that sustains the opposition between nature and nurture. The notion of information as a difference that makes a difference was then introduced. In this sense information is relational and constituted in interaction as developmental resources become relevant to the system, and it cannot be located in the genes or other components independently of the actual constructive processes in which they play their roles.

This chapter has served several purposes relevant to the focus of the thesis. If successful it should have (1) clarified some aspects of information talk in biology; (2) identified the link between ordinary talk about 'information' in the genes and, on the one hand, a view of life associated with the nature-nurture disjunction, and, on the other hand, a hylomorphic, preformationist way of thinking about development and behaviour; and (3) cleared the ground for the following chapters in which I examine ontogeny, behaviour, and social learning.

Chapter 3 The flow of ontogeny

3.1 Introduction

This chapter has three main goals. The first goal is to present a view of ontogeny – by which I mean the entire lifespan of an organism – as a fundamentally historical, relational, constructive, and contingent process. Here the organism is not conceived of as the expression of a genotype and is not equated with a specific developmental stage in ontogeny (say, the adult). Rather, the organism is conceived of as a temporally extended process in which an existing organization or form, considered as the starting state of the lifespan (usually the zygote) persists under transformations forming a dynamical pattern. This pattern *is* the organism-as-process. The transformations result from the activity of many entities and the nonlinear interactions among them. This system of nonlinear relations is denoted by the concept of the developmental system. A developmental system is defined relative to the target entity under study, such as an organism, an organ, a tissue, a cell. It includes the target entity – the current result of a history of transformations – and all interactants and processes that are currently influencing or being influenced by it. The developmental system of a target entity can be operationalised in different ways depending on the spatiotemporal resolution in which observations are made and on the questions of interest.

Given the topic of this thesis, I will pragmatically distinguish among metabolic processes, physiological processes, and behaviour, and the second goal of this chapter is to clarify the relation among them. These processes are understood as different ways to make figure-ground distinctions of what is, in fact, a single multiscale process flowing in time – namely ontogeny. The third goal is to note some consequences that follow from

the processual, developmental systems view of ontogeny that are relevant for this thesis. Thus, in this chapter I intend to clarify the basic framework that will inform the subsequent discussions about behaviour (Chapter 4) and social learning and culture (Chapter 5).

3.2 A processual view of life

I adhere to the ontological thesis that organisms are composed of a hierarchy of processes, not a hierarchy of substances. This processual view of life has its roots deeply embedded in empirical findings incorporated as part of the received view of biology, including the notions of metabolic turnover, tissue and organ regeneration, the sequence of changes that compose the typical lifespan of a species, all sorts of biological 'processes' in the usual sense, and, of course, evolution. For a recent manifesto for a process view of living systems, including a historical overview within biology, see Dupre and Nicholson (in press).

Probably the easiest way to distinguish between substances and processes is in terms of how they persist in time (DiFrisco, in press), which can be done by using the distinction between endurance and perdurance proposed by Lewis (1986). Substances *endure*, i.e. they are constituted by a set of spatial parts which do not change as time passes. Thus, in the parlance of philosophers, substances are said to be wholly present whenever they are present. On the other hand, processes *perdure*, i.e. they are composed by spatial parts that may change in time. Because processes persist by having different stages or temporal parts, they are not said to be wholly present at any moment.

One of the fundamental characteristics of organisms is that, over time, they change in some respects while remaining the same in others. This history of transformations can be observed at different spatiotemporal resolutions. At the scale of ions and molecules, for example, molecules may be taken in from the environment, change as they participate in physicochemical processes, and be eliminated to the environment. At the scale of cells, this is observed in the fascinating processes by which an initial zygote is transformed into a system of specialised cells which we call an embryo, foetus, baby, infant, and adult. At any moment, then, the organism is composed of a set of spatial parts (e.g. molecules, cells, organs, the whole body) which might change as the organism flows in time forming

a sequence of temporal parts, e.g. metabolic states, developmental stages, behavioural categories.

The typical duration in which events occur varies across processes. Thus, depending on the focus of interest, some processes can be taken as given, as unchanging, as the context, or as part of the boundary conditions in which other processes occur. For example, molecules change more quickly (metabolism) than the cells they compose, thus the global cellular organization might be taken as given and unchanging for some biochemical analysis. Similarly, cells change more quickly (e.g. metabolic states, cellular turnover) than the tissues and organs they compose, thus the global organization of the tissue or organ might be taken as unchanging in some analysis of individual cells. This insight justifies why adopting a process or developmental systems view does not require one to focus on all the processes that are going on – just like in any empirical study, most of the processes will be left in the background. However, processes under study are here taken as dynamical, relational, constructive, and contingent.

In the last chapter I rejected the idea that biological form pre-exists in the genome as information (*sensu* instruction, specification, or source of form) and I hinted at an alternative perspective focused on form-generating rather than form-imposing processes. In the following section I present the notion of self-organization in dynamical systems as a processual mechanism of how order can arise in nature spontaneously, i.e. with no need to invoke instructions (Haken, 1977; Kauffman, 1993; Nicolis & Prigogine, 1977). Because organisms realize a specific kind of organization, namely one that recursively generates the system's components, it will be appropriate to distinguish the notion of autopoiesis or self-production (Maturana & Varela, 1973, 1980, 1987) from the notion of self-organization. I then explore two related aspects of the fundamental idea that developmental systems instantiate dynamic patterns. The first aspect is that, throughout ontogeny, multicellular organisms such as animals exhibit stabilised processes that biologists commonly categorize in a compositional hierarchy of nested spatial parts (molecules, cells, tissues, organs, whole body, multispecies communities). The second aspect is that the organism is not limited to a developmental stage (say, the sexually mature stage) but encompasses the entire lifespan or ontogeny. In any period, the spatial parts that compose the body of multicellular organisms change in some respects and

remains the same in other respects, thus the organism persists under transformations in a history in which, for example, the zygote becomes the embryo, which becomes the foetus, which becomes the infant, which becomes the adult, which becomes the elderly animal. These processes of historical transformations form a compositional hierarchy of nested temporal parts which taken together, compose the flow of ontogeny.

3.2.1 Parts and whole co-emerge dynamically in systems of nonlinear relations

In the past decades, the mathematical study of nonlinear dynamical systems and the study of open systems far from thermodynamic equilibrium have provided new ways to think about how patterns – spatial, temporal, spatiotemporal, functional patterns – can arise without the need to invoke a central organizer, but rather as a result of the system's inherent dynamics, i.e. by self-organization (Haken, 1977; Kauffman, 1993; Kelso, 1995; Prigogine & Stengers, 1984; Strogatz, 2015). For a historical account, see Keller (2008, 2009).

Here are the basic ideas. In systems composed of many parts with nonlinear relations (mutualistic feedback), parts interact locally and, as a result of the cooperative and competitive effects of these local (or micro level or lower level) interactions, global (or macro level or higher level) dynamical patterns may appear in the distributed whole. The whole may become stabilised for some time and acquire a thing-like character, but its persistence depends on the continuous activity of, and interaction among, its component parts. The dynamical patterns may instantiate properties of their own, which are not specified by the parts but are inherent in the global dynamics. Thus, we can speak of emergent global processes and the properties they instantiate as being enabled by, but not reducible to, the component local processes.

Because of these nonlinear relations, the parts suddenly become systemically associated as components of a distributed whole or a synergy, and the global dynamics *constrain* the activity of the parts. To constrain, in this context, means to reduce the degrees of freedom that would be otherwise available to the parts if they were still acting as independent entities. For example, local air flows can interact locally and spontaneously produce a distributed whole or global pattern we call tornado. The tornado, once formed, constrains the possible direction of the local air flows, which may have no

alternative but to be drawn into the global pattern (say, clockwise rather than counter-clockwise), thus the form of the tornado is continually being generated by the system of nonlinear relations. Haken (1983) calls this top-down constraint, or whole-part relationship, the ‘slaving principle’.

The reciprocal influence between parts and whole, or between local and global processes, or between lower-level and higher-level processes, is nothing like a linear relationship between antecedent and consequent, or cause and effect. Local air flows do not ‘cause’ the tornado in the same way that a billiard ball ‘causes’ another one to move by hitting it. Rather, this type of part-whole relationship is captured by Evan Thompson’s elegant expression “dynamic co-emergence”, which expresses that “part and whole co-emerge and mutually specify each other” (Thompson, 2007, p. 38).

Using relative terms such as parts and wholes, local and global, lower and higher, depend on the current focus of interest. Parts (or local, or lower-level processes) interact and form a whole systemic unity. Once stabilised, the whole (or global, or higher-level process) can act as a part and interact with other stabilised processes, thus possibly bringing about further, even ‘higher’, global processes. At all levels of such nested hierarchy, each process persists under transformations, at once dependent on the ‘lower level’ processes that compose it *and* constrained by the ‘higher level’ processes which they might in turn compose.

It is possible to describe the process of spontaneous pattern formation in terms of the attractors and repellers in the phase space of the system under study. This terminology is useful, for the purpose of this discussion, because it provides an alternative vocabulary to that of computation with representations. Each degree of freedom of a dynamical system can be represented geometrically as an axis in a multidimensional space. Every possible state that the system can be in corresponds to a combination of values of these degrees of freedom and can be therefore represented as a point in this manifold or multidimensional space. The resulting volume defined by all possible values the system might exhibit corresponds to its phase space. The history of a system’s state over continuous time can be represented by a trajectory in its phase space. Because the global dynamics of the distributed whole constrains (i.e. reduces the degrees of freedom of) the local processes of the parts and their interactions, some regions in the phase space will be

visited with higher probability than others. The points or regions in the phase space to which the system's trajectories tend to converge – which might be a fixed point, an orbit, or have some other shape in the phase space – are called 'attractors'. The points or regions in the phase space from which the trajectories tend to move away from are called 'repellers'.

The layout of the attractors and repellers, i.e. the topology of the system's phase space, represents graphically the process of spontaneous pattern formation, i.e. how the inherent global dynamics constrain the possible trajectories that the system might follow over time. Given the state of the system at some moment, and its inherent dynamics, the system will tend to behave in a certain way. In other words, given the location of the system in its phase space and the attractors/repellers surrounding it, the system is more likely to follow one set of trajectories rather than another. Note that attractors and repellers are not concrete objects colliding with the system but theoretical abstractions that represent the inherent dynamics instantiated by the system. The global dynamics of a distributed system, represented in its trajectory in the phase space, co-emerge dynamically with the real-time activity of, and interaction among, the parts that compose the system.

When the system is at or near a stable state, i.e. when its trajectory is within a 'basin of attraction', fluctuations are usually compensated by the cooperative activities of other components of the distributed whole. However, when the system is at or near critical points, perturbations can be spontaneously amplified by the local feedback processes and drive the system from one attracting region to another through a nonequilibrium phase transition. When this happens, the system exhibits a new qualitative global state after the transition. The perturbations that can trigger such phase transitions can be modelled mathematically by nonspecific control parameters. They are nonspecific in the sense that they do not specify or instruct what the new state is, they do not carry information (*sensu* instruction, specification, or source of form) about the characteristics of the new global dynamics the system instantiates. The new qualitative state – the form – self-organises in the sense that it appears spontaneously with no central commander specifying or instructing it.

Take the simple case of the Rayleigh-Bénard instability. A thin layer of liquid is heated from below and kept at a constant temperature from above, such that a temperature gradient is formed. Because portions of the liquid in the lower layer become warmer and less dense they tend to rise, forming tiny transient convection streams, but this tendency is damped by the molecules bouncing around it locally. When the system is in this dynamical regime or attractor, heat is dissipated through conduction and no macromotion is observed. However, as the temperature gradient is increased, the system reaches a point in which such fluctuations are no longer damped. At some point, one of these transient convection streams eventually becomes amplified by the local interaction among molecules and the entire system self-organises a new qualitative state or attractor – the liquid begins to move as a coordinated whole in a new dynamic regime called convection. Because the macromotion of the convection cells help dissipate the heat, they are called dissipative structures (Prigogine & Nicolis, 1967). In this case, the control parameter is the temperature gradient and it should be clear that there is nothing about it that could be said to specify or instruct the form of the spatiotemporal pattern that is formed (i.e. the convection cells). The concept of a trajectory in a multidimensional phase space might be useful when discussing the histories of different dynamic systems such as a single neuron, a neuronal assembly, a region in the brain, as well as different aspects of the ontogenetic history of the coupled animal-environment system.

I did not mention organisms so far because the formation of dynamical patterns through self-organization is a phenomenon observed more generally in open physical and chemical systems that persist far from thermodynamic equilibrium. However, since the very beginning, it was clear that the study of self-organization was fundamental to understand morphogenetic processes in organisms (Eigen, 1971; Haken, 1977; Keller, 2008, 2009; Prigogine & Stengers, 1984). The notion of dynamic-co-emergence of parts and the whole they compose, or between local and global processes, will be fundamental in my discussion of behaviour, social learning, and culture.

3.2.2 Dynamic co-emergence of parts and whole in organisms

Organisms can be conceived of as nonlinear dynamical systems, in the sense that they are composed of a huge number of heterogeneous molecules which interact locally influencing each other in a continuous history of physicochemical processes (Kauffman,

1993). From this perspective, a cell is conceived of as a distributed whole or dynamical pattern that appears as a result of the local interactions among the molecules that composed it. In turn, the global dynamics that we call 'the cell' constrain the molecular interactions that might take place locally. Cells persist even as individual molecules are incorporated from the outside, transformed in chemical reactions, and eliminated. Thus, the cell and the molecules that compose it co-emerge dynamically. In the case of a multicellular organism, the nonlinear relations among molecules throughout ontogeny compose a body, i.e. a distributed system of interacting cells. The multicellular body, in turn, constrains the local cellular processes, and the molecular processes that compose it. Later we will bring the environment explicitly into the picture.

Rather than considering the whole body, we can also focus on some portion of the intricate network of (local) molecular processes which compose it. For example, within cells, a set of molecular components including phospholipids, structural proteins, enzymes, ATP, and so on, interact locally and form a global pattern we call a membrane vesicle. The vesicle, once formed, constrains further molecular interactions among its components. Individual molecules in the membrane will not be able to move around in all directions as they might if they were not caught up in the distributed whole. Rather, their hydrophobic portions are likely to remain embedded within the membrane, and their hydrophilic portions are likely to face the surrounding water, even in the absence of covalent bonds keeping them in place.

This systematic behaviour of membrane molecules is neither specified nor instructed but arises spontaneously, by means of the local interactions occurring under the constraints instantiated by the structured whole they help create. In other words, vesicles and other membrane-based cellular structures persist, not as static, substance-like things, but as perduring dynamical patterns. Note that the vesicle now separates an inside from an outside, since hydrophilic molecules in the surround are unlikely to pass from one side to another. This, in turn, influences other molecular processes in the cell because it changes the prior probability of certain chemical encounters.

Some examples of studies in biology from the theoretical perspective of self-organization processes include proteins folding (Gerstman & Chapagain, 2005), embryogenesis (Hernandez-Hernandez, Niklas, Newman, & Benitez, 2012; Newman &

Bhat, 2009), coordination of neurophysiological and behavioural activity (Di Paolo & Thompson, 2014; Haken, Kelso, & Bunz, 1985; Kelso, 1995; Schoner & Kelso, 1988; Varela, Lachaux, Rodriguez, & Martinerie, 2001; Varela et al., 1991), learning and cognitive development (Spencer, Perone, & Buss, 2011; Thelen & Smith, 1994), behavioural coordination in schooling and the formation of trails in army ants (Camazine, 2001), and interpersonal coordination (Dale, Fusaroli, Duran, & Richardson, 2013; Kelso, 1995; M. J. Richardson, Dale, & Marsh, 2014).

In sum, living systems can be viewed as intricate dissipative structures and, in that sense, they are like tornados and eddies (Nicolis & Prigogine, 1977). But organisms are not *just* like tornados. Living systems are distinguished from other complex systems because they instantiate a specific kind of organization or relationship among their component processes.

3.2.3 The autopoietic organization of organisms

According to the theory of autopoiesis, what distinguishes living from other complex systems is that the network of chemical transformations going on within its boundaries leads, among many other things, to the recursive production of its own internal components and its boundaries (Maturana & Varela, 1973, 1980; Zeleny, 1981). Through their metabolic activity, organisms continually take part in regenerating the very conditions of their persistence, which involves exchanging matter and energy with the environment. When those conditions fail to meet, the system loses its coherence and the organism dies. This insight was developed by several authors. I will mainly follow the concept of autopoiesis developed by Chilean biologists Humberto Maturana and Francisco Varela (Maturana & Varela, 1973, 1980, 1992). Other researchers have proposed similar ideas (Eigen, 1971; Gánti, 2003; Mossio, Montévil, & Longo, 2016) and it is currently an exciting area of research.

To quote one definition, autopoietic systems are "networks of productions of components that (1) recursively, through their interactions, generate and realize the network that produces them; and (2) constitute, in the space in which they exist, the boundaries of this network as components that participate in the realization of the network." (Maturana, 1981, p. 21). The autopoietic organization persists under

continuous structural changes and can be realized in different ways. Many other metabolic processes can occur and even become stable or recurrent features of an organism and the lineage of which it is a part, as long as those processes do not disrupt its autopoietic organization.

Differently from machines and other non-living systems, then, organisms are self-producing or self-making. It is worth making explicit, at this point, that the expressions self-producing and self-making are not synonyms for self-organization. The former two refer to the theory of autopoiesis, while the latter to the spontaneous pattern formation in systems far from equilibrium. The persistence of organisms depends on the recursive generation of the processes of production – denoted by autopoiesis, self-production, self-making – and this involves many instances of dynamic pattern formation through self-organization.

The use of the term ‘self’ in these expressions might suggest a primacy of ‘internal’ versus ‘external’ processes in ontogeny, or a primacy of the skin separating them as a fundamental boundary. This is not what I am implying. Ontogeny necessarily involves processes on either side of the organism-environment boundary, as well as processes that cut through this fluid boundary such as ingesting food, excreting waste, behaving, and communicating. The skin that distinguishes the body from its environment is not taken as a fundamental boundary either, but rather as one among other possibilities to draw distinctions within the organism-environment system given the focus of interest (Laland, Odling-Smee, & Turner, 2014). In the present approach, there is no such thing as a source of intelligence or form internal to the organism. There is no molecular homunculus (instructions in the genome) guiding its ontogeny from within, and there is no cognitive homunculus (instructions in the brain) guiding its behaviour from within. The self – the meshwork of interwoven processes that we distinguish as an individual – denotes the persisting organism and not an agent *within it*.

3.2.4 Developmental systems and the interdependence of organism and environment

The persistence of organisms as organised wholes is achieved through the constant activity among its components, and the network of processes of production they take part

in. This requires a constant exchange of matter and energy between organisms and the physical world around them.

In other words, organisms are fundamentally *open* thermodynamic systems, rather than closed, and must do work in order to keep on living. The cell membrane exhibits selective permeability, i.e. some ions and molecules can pass from the environment into the cells, and some can pass from the cells into the environment. Because of this, living organisms are also necessarily situated in, and continuously interacting with, relevant features of the physical world. The possibility then arises that processes in either side of the boundary under consideration, as well as processes in the fuzzy boundary itself, affect each other. Especially relevant for this thesis is the fact that this leads to dynamical patterns of organism-environment relation such as in perceptually-controlled behaviour, ecological relations, or social relations (understood as a subset of ecological relations).

Thus, the persistence of organisms depends on their surroundings providing the appropriate conditions – from the appropriate molecules to the appropriate social and ecological relations – or on the organism actively choosing or constructing those conditions (Odling-Smee et al., 2003). This makes ontogeny a fundamentally relational process. The moving, changing set of features of the physical world that are relevant for the ontogeny of the target organism corresponds to that organism's environment (J. J. Gibson, 1979/2015; Lewontin, 1983; Oyama, 1985/2000). The system composed by the changing organism and its also changing environment correspond to that organism's developmental system (Gottlieb, 2007; P. E. Griffiths & Gray, 1994; Oyama, 1985/2000). In other words, the organism, conceived of as an autopoietic system, is continually being generated as the nexus or the focal entity of an entire developmental system flowing in time.

The concept of the autopoietic system is an attempt to define what it means to be alive. Therefore, if one accepts this definition, any subsequent statement about organisms and their behaviour should be consistent with its being an autopoietic system. Organisms can only maintain their autopoietic organization as long as they are situated within and co-construct an appropriate environment, thus as part of a changing developmental system. This suggests that developmental systems theory and the theory of autopoiesis can be aligned in a harmonious way.

On the one hand, the abstract notion of a developmental system captures the empirical evidence that living systems are continually produced within a larger field of (changing) relations among components on either side of their fluid boundaries. On the other hand, researchers must always restrict their attention and define the spatial and temporal scales at which observations will be made; the specific questions being addressed as opposed to what is kept in the background; the methods employed to collect and analyse data; and so on. The abstract notion might be useful in guiding scientific practice, and it is a necessary part of any empirical investigation to find out what the relevant components of the developmental system are that make a difference in explaining the phenomenon of interest.

I will now explore two complementary aspects related to the view of organisms as processes which instantiate dynamic patterns. The first is the stabilisation of processes into thing-like entities which form a hierarchical compositional organization. The second aspect is that these spatial parts are not static but persist within histories of transformations.

3.2.5 Molecules, cells, tissues, organs, and the whole body form a nested hierarchy of composition

It is an established notion in biology that the body of a multicellular organism such as an animal is composed of organs, which in turn are composed of tissues, which in turn are composed of cells, which in turn are composed of molecules. The spatial relationships that hold among these individualized parts (say, organs) are not like the spatial relationship among the parts of a machine. Parts of a constructed machine are juxtaposed, maybe even glued or welded together. Their shape and relative positions are fundamentally determined by external agents – the people involved in their design and production.

In contrast, the different parts that compose the body of an animal at some moment do not originate independently from each other. When dissecting an organ, such as a bone, we have to quite literally cut it out from other organs and tissues to which it connects. Observed under the microscope, the tissues that compose an individual bone would reveal to be continuous with, not simply attached to, the tissues that form tendons and muscles, blood vessels, the nervous system, and so on. Organisms, in sum, are not assembled by

having parts put together. Rather, organisms grow and differentiate from previous forms within the seamless flow of ontogeny, a historical process which necessarily occurs within an environment and as part of a lineage of genealogically related organisms.

The notion of "dynamic co-emergence" is useful here to make sense of the part-whole relationships present in this hierarchical composition (Thompson, 2007). Take, for example, the relationship between molecules (parts) and a cell (the whole they compose). Molecules are not *in* the cell they compose in the same way that my socks are *in* the top drawer which contains it. Rather, the (global) process we distinguish as an individual cell *just is* the phenomenal manifestation of the dynamical patterns produced by the cooperation among metabolic processes. Once the cellular organization is established, it constrains (reduces the degrees of freedom of) the chemical processes that can occur among the molecules that compose it. That is, some chemical reactions are more likely to take place than others because of the global patterns that are already established. In other words, the cell (whole) and the entire metabolic network of chemical processes that compose it (parts) co-emerge dynamically.

Another important aspect is that the persistence of the cell does not depend on the persistence of the individual molecules that compose it at any point in time. This is because the relationship between the cell and the molecules that compose it is not one of identity but one of composition. Molecules come and go and may be transformed as they become part of chemical relations. The cell persists under metabolic turnover, and in fact depends on it, as long as the global organization persists.

In the case of multicellular organisms, individual cells (and individual tissues and organs) are not independent from each other but rather form a systemic whole. In this case, it is coherent to say that the network of molecular processes and the cells they compose, the network of cellular processes and the tissues they compose, the network of tissues and the organs they compose, the network of organs and the whole body they compose – all of these interdependent processes co-emerge dynamically as aspects of the same ontogeny. Of course, the lifespan of a multicellular organism might start with a zygote (a single cell) in its environment, and for some time the developing organism might have no tissues, no organs, and indeed no 'body'. These higher levels of composition are not there from the beginning of the lifespan but become stabilised at

different temporal stages of the ontogeny. Nevertheless, this does not mean that the molecular processes going on in these initial stages are unaffected by processes going on at 'higher' levels. This is because processes at these higher levels (say, whole bodies in their environment) are composed of molecular processes too. I will return to this point below. I should also mention that this hierarchy of composition is not bounded at the level of whole bodies. We might even go further 'up' since organisms compose populations, populations compose multispecies communities, and so on.

What I want to focus on now is the fundamentally historical and relational characters of organisms. Molecules, cells, tissues, organs, and the whole body, not only form a compositional hierarchy but co-emerge *in time* and *within an environment*. Given my interest in behaviour and learning, in the next section I focus on metabolic processes, physiological processes, and behaviour. These processes, I suggest, are three different ways of making figure-ground distinctions of the unitary, seamless flow of ontogeny, in which organisms and their environment are continually being produced. This is relevant for this thesis because it paves the way to conceive of social learning in a way that avoids the metaphor of 'transmission of information' between individuals.

3.2.6 The relationship among metabolic, physiological, and behavioural processes

The processes that compose organism-environment systems unfold at different timescales. In other words, events occurring within different processes can have typically different rates and durations. This idea is commonly expressed in the distinction between metabolism (higher rate, shorter duration of molecular processes) and development (lower rate, longer duration of morphogenetic processes), or between individual ontogenies (one lifespan) and evolutionary lineages (multiple lifespans).

The flow of metabolic processes refers to the astronomic number of physicochemical relations going on among molecules and similarly-sized components including ions, electrons, and photons. Depending on the question being addressed, relevant components might be located on either side of the boundaries that roughly separate a cell from its local surroundings, or the whole organism from its environment. The very production of the boundaries, as well as the transport of substances and particles in either direction, might be of interest too. Examples of environmental components that

take part in metabolic processes include food particles, hormones, pheromones, enzymes, photons, odorants, molecules and ions present in the medium (water, air, soil) and on surfaces of the relevant materials with which the organism interacts.

The flow of physiological processes involves the activities of distributed wholes (parts of the organism) which lead to the dynamic stabilisation of spatial structures such as organs, tissues, cells, organelles, and functional structures such as metabolic pathways. They also involve the functioning of these stabilised structures which tends to maintain the body within appropriate ranges in relevant parameters given the environmental situation. In animals, this includes processes such as digesting food, changes in the global activity of neuronal networks, moving muscles and the body parts attached to them, the circadian rhythms, regulating body temperature and heart rate. The processes denoted by physiological processes are composed of the processes denoted by metabolic processes and in turn constrain them. The relationship between metabolic and physiological processes is here seen, not in terms of linear cause and effect relations, but of dynamic co-emergence of parts and the whole they compose.

Behaviour implies a living, active animal going about engaging with aspects of its environment, as denoted by terms such as walking, grasping, and eating. Instances of behaviour are commonly described in terms of the movement of body parts relative to each other and relative to relevant features of the environmental situation (Drummond, 1981; J. J. Gibson, 1979/2015; Hogan, 2017). The animal might also be engaging with aspects which are not immediately present to the senses, such as in planning, remembering, or imagining. As Drummond (1981, pp. 3-4) summarises, "ethologists and behaviorists, though frequently divided on explanatory principles, have always recognized implicitly or explicitly that the term [overt behaviour] encompasses all observable activities of an organism" thus highlighting "the integration of organism and environment" that underlies the "stream of behavior". This general characterisation of behaviour is consistent with the approach to social learning that I am beginning to develop in this thesis.

However, cognition is commonly described from a computational perspective, and the form of some observed behaviour is commonly explained by pointing to a previous instance of that form. Similar to how 'information' encoded in the genome is commonly

assumed to be the source of biological form, 'information' encoded in the nervous system (as a program, representation, algorithm, and the like) is commonly assumed to be the source of behavioural form. These claims express a prior commitment to a theoretical view about what cognition is about and how behavioural control is to be understood. It is, however, a theoretical choice, and not a necessity, to describe cognition in such terms. My theoretical choice, in contrast, is to start with the idea that cognitive events unfold in the animal's real time coupling with the environment rather than in the runtime evaluation of computational algorithms.

What, then, one might ask, is the relationship between the activity of the nervous system and the flow of behaviour? The flow of neurophysiological activity in the nervous system is, of course, one aspect of the animal-environment system relevant to understanding how behaviour is produced, and the role of the nervous system in coordinating behaviour can and should be studied. The critical point, though, is that neurophysiological processes refer only to a part of the system of nonlinear relations involved in cognition and behaviour. Rather than adhering to the representational-computational approach to cognition, my goal is to investigate behaviour and social learning from the perspective of radical embodiment.

The stream of behaviour and the stream of physiology flow in time as aspects of the same ontogeny. The movement of a body part during, say, extending a leg, occurs during the exact same period as the (physiological) changes in the contracting state of the relevant muscle fibres and the exact same period as the (neuro)physiological changes in the activity of the relevant neuromuscular networks (Kelso, 1995; Thelen & Smith, 1994). Talking about behaviour, metabolism, or physiology, are different ways to make figure-ground distinctions relative to the ontogeny under study. In discussing behaviour, a 'big picture' description of the activity of the whole body-in-the-environment – with a coarser spatial resolution and with a special interest in what is going on between the surface of the animal and its environment – is brought to the foreground of the analysis while other aspects of the ontogeny are kept in the background. On the other hand, in discussing (neuro)physiology, a small set of the activity of the nervous system is brought to the foreground instead. In discussing metabolic processes, a small set of physicochemical relations are brought to the foreground. What distinguishes the flow of behaviour from

the flow of physiological adjustments is mainly the focus of interest of the observer and the spatial scale of the relevant components being invoked.

This view of the relation among these related processes is preferable, I argue, because it does not mistake changing scales and focus of interest for going from causes to effects. It will be noted that this suggestion is not an answer to the question, "What, then, causes behaviour if not the activity of the nervous system?". I will return to this question in the next chapter.

3.3 Consequences of a processual view of life

In the last part of the chapter I focus on some of the consequences that follow from the processual, relational-historical, or developmental systems view of ontogeny presented above.

3.3.1 Development is a possible consequence of ontogenetic processes

Development refers to morphogenesis, i.e. changes in form (including morphological, physiological, and behavioural form) that organisms undergo throughout their lifespan (Gilbert, 2010). Development might refer to changes in entities and processes observed at any spatiotemporal resolution, including changes in the molecular composition of cells, in the size and state of tissues and organs, in the abilities to perceive and act, and so on. The relatively stable patterns that appear in ontogeny are commonly denoted as developmental stages.

Rather than describing major changes in the historical process of zygote-in-its-environment becoming body-in-its-environment, we might narrow our focus on some specific aspect of interest – the target trait in the study – and investigate the processes that compose its history of becoming. For example, we could trace and investigate the temporal parts or developmental stages of a bone such as a femur from the embryonic tissues, or the embryonic stem cells, or the initial zygote. We could also trace and investigate the temporal parts or developmental stages of abilities underlying behavioural patterns (Spencer et al., 2011; Thelen & Smith, 1994). Learning, therefore, might be considered a class of developmental changes in the bodily structures underlying behaviour (Hogan, 2017). In dynamic parlance, learning can be conceived of as changes in the layout of attractors in the behavioural phase space defined by the animal-

environment system (Schoner, Zanone, & Kelso, 1992; Smith & Thelen, 2003; Thelen & Smith, 1994).

Developmental changes appear as *consequences* of certain causal relations among relevant components of the organism-environment system, and they can be triggered by entities located on either side of the boundary. Whatever developmental effect an entity may have – a hormone, a nutrient, a DNA segment, a volatile substance dissolved in the olfactory mucus, a pattern of neuronal firing as a result of behaviour, and so on – depends not only on the structure of the entity itself but on the relation it instantiates with the part of the organism it affects, the current state of the organism, the timing of the influence, and so on. Activating 'the same' gene can have different effects in different tissues and in the same tissue at different times. Perceiving 'the same' environmental situation – including the presence and behaviour of facilitators in the context of social learning, can have different effects depending on the state of the organism and its history of behavioural events thus far.

3.3.2 Ontogenies are nested within evolutionary lineages

The events of interest in this chapter occur within the period of individual lifespans, but we should note that ontogenies are nested within higher-order processes. A collection of ontogenies of the same species living together compose a population, and usually populations form multispecies communities with mutual involvement, as their individual ontogenies become woven together in time by ecological, social, and genealogical relations. Any individual ontogeny begins as a transformation of some part (e.g. gametes) of one or two other, parental ontogenies. Individual ontogenies interact locally and the global spatiotemporal patterns they produce constrain (reduce the degrees of freedom of) the historical trajectories of individual ontogenies. In other words, individual ontogenies and multispecies communities co-emerge dynamically. The resulting evolutionary flow is therefore not orthogonal to the ontogenetic flows, but composed of them, and unfolds at a larger timescale.

3.3.3 Ontogeny is a fundamentally historical, relational, constructive, and contingent process.

One of the starting points of this chapter was a commitment to process ontology and developmental systems thinking. One of the fundamental characteristics of processes is

that they persist over time as a sequence of temporal parts, rather than by being numerically identical. Organisms are never finished or fixed entities but undergo continuous transformations as they keep on living. Ontogeny, as used here, refers to the sequence of temporal parts that compose the entire lifespan of an organism. In the case of most animals, ontogeny begins with a zygote and ends only when the organism dies. Ontogeny is therefore a *historical* or dynamical process.

Ontogeny is a *relational* process because this history of transformations results from the real-time interaction among the many components of the organism's developmental system (Gottlieb, 2007; P. E. Griffiths & Gray, 1994; Oyama, 1985/2000). These components might be detected and described by us observers at different scales of size, and they might be located on either side of the fuzzy boundary between an organism and its environment. They include non-living as well as living entities. Therefore, an organism and its environment are interdependent aspects of the same ontogenetic process and one cannot be defined without implying the other. The term developmental system denotes the moving set of components defined by the changing organism and its changing environment.

Ontogeny is a *constructive* (or emergent) process not only because many of the components which take part in this relational history of transformations are constituted in the process, but because the (changing) form of the whole organism and the (changing) form of its environment are themselves constituted in the very process. Ontogeny is *contingent* because there is no script laid out in advance for this history of transformations. The ontogeny trajectory of any organism is enabled and constrained by the available developmental resources – some reliably present across generations, others less so – but these resources do not specify the emergent form from the outset. The availability of similar resources – e.g. because they are recurrent or persistent features of the physical world, or because they are somehow passed on between organisms – may lead different organisms to undergo similar ontogenetic trajectories, but it would be a logical mistake to conclude that any of those resources contain something like instructions or a script for ontogeny (Kay, 1998; Keller, 2000; Oyama, 1985/2000; Sterelny & Griffiths, 1999).

3.3.4 Ontogeny unfolds from phenotype to phenotype

Ontogeny is commonly said to be a process that goes from genotype (or genes, or the genome) to phenotype, but this way of speaking is misleading. It is conventional biological knowledge that genotypes, or genomes, are never found except as a part of an already organised organism and cannot make things happen in the organism except as constrained by a system of relations with other molecules which they causally interact. It is also a robust empirical fact that even zygotes are much more than just a collection of DNA molecules ready to start constructing the future organism out of inert raw materials – indeed, the zygote is the *initial* form of the organism, i.e. a temporal part of the organism-as-process. The initial molecular composition of the zygote –including its DNA and much more – plays a fundamental, constructive role in development. And if zygotes are to keep on living, the chemical environments around them must also exhibit an appropriately changing composition and organization.

Rather than saying that an ontogenetic trajectory goes from genotype to phenotype, then, it would be more precise to say it goes from phenotype to phenotype – or, to be more precise, from an early and younger version of the phenotype (with all that comes with it, including of course its DNA) in its environment (with all that comes with it, including other organisms and their stabilised patterns of relating to each other) to a later and older version of the now transformed phenotype in its also transformed environment.

3.3.5 The social zygote

The initial temporal parts of the organism (e.g. the zygote, the blastula, and the morula) have no organised tissues or organs, but that does not mean that ontogeny, at these early stages, is not constrained by what is going on at 'higher' levels of organization. In the case of placentary mammals, for example, the zygote is formed within the reproductive system of the female parent. The parental organism is not an abstract bucket of genes but a material, living system already organised into cells, tissues, organs, and a body. Parents are already actively engaging in *their* environment. Therefore, the activities of parental organisms co-determine the molecular composition inside and outside of the zygote, thus constraining the possible ontogenetic trajectories available to the zygote-becoming-embryo-becoming-baby. While their offspring are developing, the caregiving parents are still busy living their lives, their own ontogenies unfolding in an ongoing history of

interactions with other organisms in the social group and the multispecies ecological community of which they are a part (Ingold, 2000; Toren, 2012).

Therefore, the patterns of relations that are already established in a social or multispecies community, at the same time, both enable and constrain the ontogenetic trajectories of the arriving organisms. The notion of dynamically co-emerging ontogenetic flows offers a way to understand how stabilised spatiotemporal patterns of social relations, even those denoted by the term social institutions, can play relevant, constructive roles in codetermining the developmental trajectories of a person or nonhuman animal. They do so not as a reified, "superorganic" (Kroeber, 1917) thing – society or culture – acting on the animal from the outside but by co-determining the components and processes that can become part of its developmental system, and therefore co-determining the interactions that are effectively available to the living animal. This is the case even if, at the youngest stages, the newcoming animal has very limited possibilities to engage with old-timers in the communities of which it is now a part.

3.3.6 No place for genetic programs

When the development of a feature is robust (i.e. its development is not affected by the manipulated developmental conditions), similar across organisms, and stable across generations, it may give the impression of being programmed, of being directed by an internal homunculus. The notion of genetic programs has been widely criticised (Keller, 2000; Neumann-Held & Rehmann-Sutter, 2006; Oyama, 1985/2000, 2015; Oyama, Griffiths, & Gray, 2001; Sterelny, 2000; Sterelny & Griffiths, 1999) and it is here assumed to provide an inadequate account of morphogenetic processes. Problems involving this notion include its preformationist logic; how it helps the "mirage of a space between nature and nurture" (Keller, 2010) to persist; how it is linked with a failure to distinguish between sources of variation and sources of form (Keller, 2010; Oyama, 2000a); and how it is linked to an ignorance of not-so-recent-anymore advances in the formal study of complex systems, which present an alternative view of how form may arise out of disorder in a system of nonlinear interactions (Haken, 1977; Hernandez-Hernandez et al., 2012; Kauffman, 1993; Kelso, 1995; Newman & Bhat, 2008; Prigogine & Nicolis, 1967; Prigogine & Stengers, 1984).

However, the habit of invoking the genetic program metaphor as a seeming explanation for development is a common behavioural pattern among biologists; it seems to be robust relative to the available critiques and to be reliably recurrent since its initial development (F. Jacob & Monod, 1961; Mayr, 1961). It is indeed impressive how the development of this habit seems an almost inevitable outcome of the process in which young people become biologists. It would be an interesting question to ask which learning processes are involved in its remarkably reliable recurrence within our educational systems.

3.3.7 No place for nature-nurture

The distinction between two types of developmental processes – one to reveal programmed traits and another for everything else – has no place in the current view of development as a dynamic, relational, and constructive process. As a result, the nature-nurture conundrum, in its many forms (including open programs-closed programs, instinctive-learned, innate-acquired, biology-culture) simply does not arise.

3.3.8 Traits as histories

The empirical fact that organisms have temporal parts (as well as spatial parts) has a profound impact on the notion of biological traits. At any moment, an organism has a current form which is the result of the ontogenetic history that has brought it thus far. Like the whole organism of which they are a part, traits – i.e. any observable features of the organism that are of interest, at any scale of size, including behavioural patterns – also have a history. This is necessarily the case even if we abstract a snapshot from this history in scientific practice. The term that we use as a label for, say, a trait that is dynamically stabilised in the adult animal, implies that ontogenetic history and its fundamentally relational, constructive, and contingent character. Labels for traits, then, are like condensed narratives of the contingent history of interacting processes which made them what they are at the time of observation.

3.4 Summary and conclusion

Organisms are open systems far from thermodynamic equilibrium. To say that organisms are open systems rather than closed ones refers to the constant exchange of matter and energy with the environment. Thus, organisms and their environment are interdependent

aspects of the same system – the developmental system. To say that organisms are systems far-from-thermodynamic equilibrium, rather than at or close to it, refers to the ability of organisms to persist as incredibly organised wholes even in the face of the universal tendency of systems to disintegrate and to lose their internal organization. Organisms, that is, perform work (in the physical sense) as they keep on living.

What distinguishes a living from a non-living system is that the network of chemical transformations going on within its boundaries leads, among many other things, to the recursive generation of its own internal components and its boundaries. Thus, cellular metabolism realizes a specific molecular organization which entails the continuous generation of itself. This requires that organisms find and, in many cases, participate in creating suitable environments in which to live.

In systems far from thermodynamic equilibrium such as organisms, when many parts constrain each other's activity, these (local or micro-level) relations may lead the system to spontaneously organize (global or macro-level) spatial, temporal, or spatiotemporal patterns which instantiate new properties. These new or emerging properties are measured by variables which distinguish those aspects of the whole which are not reducible to 'lower-level' variables describing the components. In turn, once these global patterns are established, they constrain (i.e. reduce the degrees of freedom of) the local parts. In this process of spontaneous pattern formation, the relation between parts and whole are not captured by the notion of linear causality. Instead, this relationship is better captured by the notion of dynamic co-emergence. This notion was invoked to make sense of part-whole relations in organisms.

The individual ideas presented above are not new. Arguably, the originality of this chapter lies in the way in which I consistently weave them together to highlight their relation and clarify points of potential conflicts. I started this chapter with a processual and relational view of organisms. The spatial form of organisms at any moment was here conceived of, not as a nested hierarchy of juxtaposed, independent substance-like things, but rather as a nested hierarchy of stabilised processes that co-emerge dynamically. And the history of transformations that organisms undergo in their environments throughout their lifespan was conceived of, not as a sequence of juxtaposed, independent states, but rather as temporal parts of a temporal whole. I distinguished aspects of the flow of

ontogeny – metabolic processes, physiological processes, and behaviour – that will be relevant for the topic of social learning. Within this view, an individual ontogeny unfolds as a network of fundamentally relational, constructive, and contingent processes, going on at different scales of size and time. I ended by noting some consequences that follow from this view of ontogeny. In the next chapter, I expand on the topic of behaviour before finally turning to social learning and culture.

Chapter 4 The flow of behaviour

4.1 Introduction

In the previous chapter I presented a view of ontogeny as a dynamical, relational, constructive, and contingent process, and I distinguished behaviour as one aspect of the flow of ontogeny. In the parlance of dynamical systems, the flow of behaviour is a process of dynamic pattern formation. To exhibit a behavioural category is to enact a trajectory in the phase space defined by the organism-environment system. To describe overt behaviour is to describe this process from a certain viewpoint (as external observers), with reference to a specific scale of spatiotemporal resolution (e.g. distinguishing units such as limbs and objects rather than individual cells and molecules) and with a specific epistemic interest (e.g. the movement of body parts in relation to each other and in relation to features of the environment, rather than cellular metabolism).

To say that behaviour is a trajectory in phase space is admittedly vague. It is, I suggest, just as vague as saying that behaviour is the output of computations with representations. These suggestions are, in a way, meant to be vague, since they express a general metaphor or starting assumption about what behaviour is. One difference between these two suggestions might be that we behavioural scientists are likely to be much more familiar with computers, with the common-sense notion of 'processing information', and with the idea that we are centralized agencies controlling our bodies from the inside. It is therefore not surprising that the computational metaphor might feel more familiar than the dynamic system metaphor. However, it is not familiarity, but harmony with a relational-processual approach, that guides my preference for the latter. By making this methodological choice, my intention is to highlight the mutual support among developmental systems perspective (Oyama, 1985/2000; Oyama et al., 2001), radical

embodiment (Chemero, 2009; J. J. Gibson, 1979/2015; Varela et al., 1991), and relational thinking (Ingold, 2000, 2011a; Ingold & Pálsson, 2013; Lave & Wenger, 1991; Toren, 2008, 2012).

Conceptually, then, I view behaviour as a dynamic pattern produced by the animal-in-its-environment, rather than as the expression of the activity of the nervous system. This should *not* be taken as a suggestion that the nervous system plays no role in producing behaviour, but as an acknowledgement that the activity of the nervous system not only influences but is influenced by the activity of other components. Some components are located inside the body, e. g. bones, joints, muscles. Others are located outside the body (e.g. rocks, water bodies, other organisms). Others are found in the fuzzy boundaries between the animal and its environment (e.g. photons interacting with the retina, volatile molecules interacting with the olfactory epithelium). The mutual constraints among these components reduce the degrees of freedom that would otherwise be available to each if they were not part of a distributed whole. Neurons could be firing at a different rate if they were not interacting with other neurons and with sensory stimulation. Skeletal muscles could be in another state of contraction if they were not connected to neurons and bones. Bones could be in a different relative position if they were not pulled by muscles and constrained by joints. The movement of body parts is also constrained by the medium in which the animal lives (e.g. air or water) and the surfaces with which they interact mechanically (e.g. the ground, objects).

Because these components form a system of nonlinear relations, and because all these components influence the form of the behaviour that animals end up exhibiting, it is not possible to single out one component – the nervous system – as 'the' central controller. It is, however, certainly possible to try and distinguish the different roles played by different components and to study how each constrains, and is constrained by, the others as they produce the dynamic pattern we call behaviour.

The goals of this chapter are to (1) present a non-representational, non-computational view of behaviour that is original in its attempt to be consistent with a processual, developmental systems view of ontogeny and in its link with social learning; (2) clarify how I understand the relationship between the activity of the nervous system and other components of the developmental system with respect to behaviour, (3) suggest

how the flow of behaviour is produced (if not controlled by the nervous system), and (4) conceptualise learning as a possible consequence of behaviour. This will set the conceptual context for the discussion about social learning and culture in the remaining chapters.

4.2 What cognition is about

To achieve the goals of this chapter, I need a view of behaviour – indeed, of cognition in general – that is consistent with my premise, that is, one in which behaviour and cognition can be conceived of as thoroughly historical, relational, constructive, and contingent processes. In other words, I need a view of behaviour and cognition that takes the changing animal-in-its-environment as the fundamental unit of analysis. This will not be the dominant, representational-computational view of cognition. The cognitivist view is based on the representational theory of the mind (Fodor, 1975). It starts with the theoretical suggestion that cognition involves the operation of mental processes on mental states. Mental processes refer to activities such as thinking, remembering, comparing, judging, and inferring. Mental states refer to thoughts, beliefs, desires, perceptions, memories, and so on. These states are said to have intentionality because they are *about* or *refer to* things other than the states themselves. For example, the perception of a red apple may be about a fruit that exists in the world, and a memory state might be about an event which occurred some time ago. The representational view of cognition explains intentional states in terms of representations, and explains their intentionality in terms of the semantic properties of representations (Pitt, 2017).

In its contemporary, computational version, mental representations are taken to be analogous to computer data structures, and mental operations are taken to be analogous to computational algorithms (Thagard, 2014). Mental representations, like data structures in the computer, are considered to be information-bearing structures (Pitt, 2017). That is, they carry information (*sensu* what-it-is-we-can-lean-from-it) about that which they represent. For this reason, mental or cognitive processes are, like computational processes, commonly characterised in terms of the 'processing of information' (although it would be more coherent to say processing of data). Most behavioural biologists work from this perspective and commonly conceive of learning in terms of the acquisition of task-relevant 'information', and social learning as a process of 'transmission of

information' from one animal to another (Boyd & Richerson, 1985; Mesoudi, 2011; Whiten, Horner, Litchfield, & Marshall-Pescini, 2004). In these cases, the term 'information' is used somewhat loosely to mean, in fact, 'representation', possibly as a shorthand for 'information-bearing structure'. It is therefore intimately associated with the notions of (mental) representation and computation.

4.2.1 Changing the basic metaphor: from computation to dynamics

In this thesis I reject representational-computational views of cognition as a starting point. However, my intention in the rest of this section is not to argue against it but simply to contrast it with an alternative which is non-representational and non-computational and therefore consistent with the approach I am developing.

In a paper titled "What might cognition be, if not computation?" Tim Van Gelder (1995) famously discussed the operation of the governor of steam engines, developed by Scottish engineer James Watt, to argue that cognition can be understood in dynamical rather than computational terms (see also a relevant discussion of this paper in Chemero, 2009, chapter 4). The answer Van Gelder gives to the question in his title is the following. "Rather than computers, cognitive systems may be dynamical systems; rather than computation, cognitive processes may be state-space evolution within these very different kinds of systems." (p. 347).

I will follow this suggestion and side with the alternative view known as radical embodied cognitive science. I use this term after Chemero (2009) to refer to two traditions. One is the ecological psychology developed by James J. Gibson and Eleanor J. Gibson, and followers (E. J. Gibson & Pick, 2000; J. J. Gibson, 1966, 1979/2015; Kelso, 1995; Lee, 2009; Turvey, Shaw, Reed, & Mace, 1981) The other is the enactive approach initially proposed by Varela et al. (1991), following the insights from the theory of autopoiesis which Varela developed with his former professor Humberto Maturana (Maturana & Varela, 1973, 1980, 1992). Although the two approaches developed as different traditions, it has been increasingly recognised that they share more similarities than differences, and that it might be possible to bring them closer together conceptually (Chemero, 2009; McGann, 2014; Thompson & Varela, 2001).

I am less interested in what sets these two traditions apart than in what makes them natural allies and for the most part I will treat them together under the banner of radical embodiment, still making distinctions where appropriate for clarity. This is how Chemero (2009, p. 160) summarises the approach:

> Here, then, is radical embodied cognitive science: animals are active perceivers of and actors in an information-rich environment, and some of the information in the environment, the information to which animals are especially attuned, is about affordances. Unified animal-environment systems are to be modelled using the tools of dynamical systems theory. There is no need to posit representations of the environment inside the animal (or computations thereupon) because animals and environments are taken, both in theory and models, to be coupled.

This is, indeed, consistent with my purpose. The use of the term information in the context of radical embodiment is different from its use in the computational approach. I will return to this point below. For now, I want to consider a few issues to illustrate how the dominant approach (computational, representational, cognitivist) and the alternative approach (non-computational, non-representational, ecological, enactive, radical embodiment) diverge. According to the cognitivist view, perception involves forming meaningful representations from the meaningless stimuli that arrive at the sensory organs by means of computational processes operating on them. In contrast, according to the radical embodiment alternative, perception occurs in the dynamics of the coupled animal-environment system and it is therefore unnecessary to posit representations of the world inside the animal. In saying this I do not mean to be providing an explanation for how perception occurs but only to clarify how the different approaches suggest different starting points for theoretical and empirical studies of perception.

In the cognitivist approach, action (overt behaviour) is taken to be the expression of motor programs, the output of rule-based manipulations of representations both obtained in perception and retrieved from memory. This implies that the form of overt behaviour – how the animal moves in space and time during a period of time – is somehow given in advance in the posited motor program prior to its appearance in actual bodily movements. The alternative perspective rejects this version of preformationist thinking. The form of behaviour cannot possibly be given in the nervous system prior to its

appearance because, in addition to the nervous system, other components – on either side of the skin – play causal roles in determining the flow of behaviour in real time.

The suggestion that behavioural form is already encoded in a previous motor program prior to its appearance in the actual behaviour is analogous to, and as problematic as, the notion that phenotypic form is already encoded in a genetic program prior to its appearance in ontogeny (Oyama, 1985/2000). From a developmental systems perspective, biological form – including behavioural form – is neither pre-existent nor ever finished. Rather, form appears and persists under transformation as the result of historical, constructive interactions going on among the components of the organism's developmental system, as ontogeny unfolds. In a nutshell, behaviour does not appear in the runtime execution of an already computed algorithm, but in the real-time engagement of the animal with its environment.

4.3 What behaviour is about

Behaviour involves perceiving relevant aspects of the physical world and acting upon them and in relation to them. It might also involve awareness of aspects of the environment that are not immediately present to the senses, such as remembering situations lived in the past, planning, or imagining. To make sense of how the flow of behaviour is produced (if not by manipulating mental representations), I will discuss the notion of 'environmental information', its relation to abilities to perceive and act, and the role of intentions. I hope to show that the notion of environmental information is quite different from 'information' (meaning representation) as used in the cognitivist approach.

4.3.1 The information available in the environment

The environment of an animal usually includes non-living components such as the medium (water for aquatic animals, air for terrestrial animals), the 'natural' objects and materials such as mountains, hills, rocks, and water bodies, as well as the objects and materials produced by the niche-constructing activities of organisms such as body fluids and waste materials, oxygen and carbon dioxide, nests, bee hives, termite mounds, shelters, and the myriad of human-made artefacts. These objects and materials furnish the different places in the environment which might provide animals with different behavioural opportunities or affordances.

Consider how the ‘information’ available for visual perception is created (J. J. Gibson, 1979/2015). In an illuminated environment, light coming from the sun or other sources of illumination is thoroughly scattered by bouncing back and forth between reflective surfaces and the medium, and it achieves an equilibrium or steady state. Usually, the light arriving at a fixed point of observation is different from different directions in a way that is specific to the surfaces surrounding that point. In other words, there is a structured rather than a homogeneous flux of photons arriving at any point of observation.

In his ground-breaking work, James Gibson conceived the structure of the optic array as a nested set of solid angles, with their apices at the point of observation and their bases at distinct surfaces. Neighbouring surfaces, or environmental texture elements, can be distinguished by the intensity and/or spectral composition of the light they reflect. The structure or pattern at the proximal point of observation specifies – is specific to, and thus can *inform about* – the distal surfaces which constrained its structure. It is in this sense that the ambient light array reaching a fixed point of observation and, especially, the systematic changes in the array produced by moving along a path of observation (the visual flowfield), can be said to create ‘environmental information’. If there was no structure in the light arriving at the point of observation, as in a room filled with white fog, there would still be stimulation of the sensory neurons but there would be no stimulus information available to be detected.

According to J. Gibson, “*Information*… refers to specification of the observer’s environment” (J. J. Gibson, 1979/2015, p. 131, emphasis in the original). It emphatically should not be taken to mean a piece of knowledge conveyed in communication, a representation, or a computer-like data structure. He was very much aware that the computer metaphor was becoming increasingly popular, with the expression information processing being used in place of the traditional notion of input processing by mental operations. J. J. Gibson (1979/2015, p. 240) argued that the move was hardly innovative:

> But it seems to me that all they [researchers using the computer metaphor in the study of perception] are doing is climbing on the latest bandwagon, the computer bandwagon, without reappraising the traditional assumption that perceiving is the [mental] processing of [meaningless, sensory] inputs. I refuse to let them

> pre-empt the term information. As I use the term, it is not something that has to be processed. The inputs of the receptors have to be processed, of course, because they in themselves do not specify anything more than the anatomical units that are triggered.

Environmental information should also be distinguished from how the term 'information' is commonly used in the context of communication technology. "The information for perception is not transmitted, does not consist of signals, and does not entail a sender and a receiver" (J. J. Gibson, 1979/2015, pp. 56-57). In another passage, Gibson explicitly distances his use of the term from the mathematical concept. "The information for perception, unhappily, cannot be defined and measured as Claude Shannon's information can be" (J. J. Gibson, 1979/2015, p. 232).

Gibson's use of the term 'information' has influenced later work by situation theorists, which I mentioned in Chapter 2. Barwise and Perry (1981, p. 668), for example, explicitly say that their view "was profoundly influenced by [ecological psychologist Michael] Turvey and others working in the tradition of ecological realism". According to situation theory, information is created by 'situations' and informational relations between situation tokens should be understood in terms of 'constraints' (lawful or normative regularities) between situation types. To use Dretske's (1981) expression, the informational content of a situation is the 'what-it-is-you-can-learn' about one part of the world (situation, signal, state of affairs) by consulting or detecting some other part of the world. Situation theorists were initially interested in linguistic phenomena. Translated to perception, we can say that the structure of the environment and the structure of the ambient array of light (as well as sound, chemicals, surfaces, and so on) are linked by regularities or constraints. Thus, the animal can learn about, or become aware of, its environment by detecting the ambient arrays that specify it.

4.3.2 Environmental information guides behaviour

When an animal occupies a point of observation, its sensory organs can be stimulated by the structured ambient arrays of light, sound, chemicals, surfaces, and so on. Moreover, its body takes part in structuring the ambient light together with the other objects and, therefore, the information available at that point is as much about the situations it involves as it is about the animal itself. That is, "the perceptual systems are propriosensitive as

well as exterosensitive" (J. J. Gibson, 1979/2015, p. 108). The vague boundaries of the field of view, for example, specify the body occupying that location – more specifically, the nose, the edges of the eye socket, the eyebrows, and the cheekbones. Other parts of the body such as arms, hands, abdomen, legs, and so on, are also directly specified in the ambient light array. Information about the animal is available to all perceptual systems: "An individual not only sees himself, he hears his footsteps and his voice, he touches the floor and his tools, and when he touches his own skin he feels both his hand and his skin at the same time. He feels his head turning, his muscles flexing, and his joints bending." (J. J. Gibson, 1979/2015, p. 108).

Suppose you look at a rectangular table by keeping one eye shut and positioning the open eye on a line perpendicular to its surface. The form of the envelope of the solid angle arriving at the retina, which specifies the outline of the face of the object as seen from this point of observation, will be rectangular (or nearly so, for the retina is not a plane). If you now move a bit to the side, the form arriving at the retina is no longer rectangular but trapezoidal. This is the case because the angles at the table corners and the proportions of the sides change from one point of observation to another. At the same time, there are relations among the angles and relations among the proportions, described by projective geometry, that remain unchanged under transformations. These invariants, as Gibson argued, are specified in the optic flowfield and available to a moving observer. The suggestion is that, by looking at the table along a path of observation, your perceptual system can detect the underlying invariants and you therefore perceive the form of the table as a persistent feature of the environment.

This perception is direct in the sense that the patterns are present in the visual flowfield and therefore it is not necessary to posit mental operation such as representing the table's angles and side lengths and then performing computations to derive the invariant quantities. Rather than computing the invariants, perceiving is about detecting them. More recently it has been suggested that pattern *detection* can be achieved, without recourse to algorithmic operations, by spontaneous (self-organising) processes of dynamic pattern *formation* in nonequilibrium systems (Haken, 1996; Kelso, 1995). In the case of visual perception, such processes may result from the perturbations produced in the ongoing neurophysiological dynamics by the causal interactions between photons and

photoreceptors in the retina. Saying this does not solve the problem of how visual perception occurs but it offers an alternative way to conceive the problem in non-representational, non-computational terms. Much more empirical and modelling work is required to understand the details of how direct perception is achieved. It should be noted that saying that the brain really computes the invariant quantities (instead of sensing it directly) does not solve the problem either, but rather defines the alternative computational framework which might be used to investigate it.

When an observer moves along a path of observation, the optic structure arriving at the retina (i.e. the environmental information thus made available to be picked up by the visual system) changes in some respects – what Gibson called the 'perspective structure' – and remains unchanged in other respects – the 'invariant structure'. Each unique point of observation in the environment is specified by a corresponding perspective structure. Each line segment connecting adjacent points of observation, i.e. each path of possible displacements in the environment, is specified by a corresponding set of transformations in the perspective structure. Therefore, the information made available in the time-evolving perspective structure, which in part depends on the activity of the animal, specifies the current position and the path of locomotion of the observer in the environment. That is, a perspective structure flowing in time *means* locomotion along a specific path, and an arrested perspective structure *means* rest. The optic structure that remains invariant under transformation requires activity to be made available and sensed. This invariant structure is common to a set of points of observation and specifies, not the observer, but the persisting features of the environment.

Events occurring in the environment are specified by local disturbances in the ambient structured arrays. For example, the displacement of an object against a background is specified by the progressive deletion of optic structure at one border and accretion of structure at the opposing border. The approach of an object is specified by *looming*, i.e. the magnification of the optic form that specifies the object with progressive deletion of optic structure outside its contour. Elastic events such as the overt behaviour of other animals are specified by the deformation of the optic form that specifies their body parts. Vocalisations structure the mechanical waves in the medium and are therefore specified in the sound array. The elimination of substances, such as releasing pheromones

in the atmosphere or leaving chemical trails on the ground, structure the available chemical arrays. And so on. It is invariants such as these that animals have available to detect the presence and activity of other animals, and which they might use to guide their own behaviour. J. J. Gibson (1979/2015) argued and reviewed empirical evidence that people are indeed able to detect invariants, and research programs inspired by Gibson have provided further support (Lee, 1976, 1998; Lee & Reddish, 1981; Turvey & Carello, 2011).

Note that the information available in the perspective structure is propriospecific, i.e. is about the relative position of the animal in the environment. In contrast, the information available in the invariant structure is exterospecific, i.e. is about the components in the environment, including other animals. They are concurrent and each imply the other. By picking up both kinds of information, animals might perceive the environment and themselves in it at the same time.

Animals in different locations, and animals following a similar path at different times, might detect the same patterns. The fact that environmental information is thus publicly available to animals with similarly tuned perceptual systems is crucial to understand processes of social learning without recourse to the metaphor of 'transmission of information' (representations) between animals. Animals that are closely related genealogically, and whose ontogenies unfolded in similar developmental niches, might develop functionally similar perceptual systems. Consequently, they might detect the same information available in their common environment.

Also crucial to social learning is that animals playing the role of facilitators can, by means of their presence and activity, direct the learners' attention to information specifying task-relevant aspects of the environment as well as create information that would otherwise not be available to them. When facilitators perform the target task, vocalize in response to a perceived threat, or provide verbal instructions, for instance, they structure the ambient arrays of light, sound, etc, thus creating information that might be perceived by others around them. In this view, social learning is a historical and relational process of developing abilities, in which ontogenies unfold in time and are intertwined, rather than a computational process of acquiring representations. I will return to this below.

In the ecological approach, perception is conceived of as a continuous activity of sampling the ambient arrays of light, sound, chemicals, materials, and so on, that specify features of the environment and the perceiver at the same time. Some of the information available in the environmental situation is especially relevant for behaviour because it specifies the affordances of surfaces, objects, places, and events. For example, if we perceive that an object has opposing faces separated by a distance smaller than our hands, we perceive that the object affords grasping. The affordance graspability is specified in the flowing optic array and can be perceived directly, i.e. without invoking algorithmic transformations of representations. To argue that affordances can be perceived directly does not imply that animals 'just do it'. The abilities required to detect environmental information and perceive affordances, like all abilities, must of course develop, and can also be changed consequent on experience, i.e. through learning (E. J. Gibson & Pick, 2000).

Animals with appropriately tuned perceptual systems can sample the patterns available in their current environmental situation and use them to adjust their flow of behaviour accordingly. This includes guiding their behaviour in relation to aspects of the environment that are currently present to their senses. For example, the optic flowfield being such and such might *inform* the animal *about* the presence of a fruiting tree ahead. The skilled animal can use the patterns in the optic flowfield to adjust the direction of its locomotion towards it. Similarly, when linguistically skilled people engage in conversation, the mechanical (sound) waves being such and such *inform* them *about* what the other wants to convey and they can adjust their own utterances in response.

Animals can also use the patterns available in their current situation to become aware of the past, present, and future situations they imply, because of the regularities linking them. Skilled animals can therefore guide their behaviour in relation to aspects of the environment that are *not* currently present. Imagine a group of hunters searching for game as they move along a path in their environment. On the one hand, the (proximal) patterns in the optic flow as they move along specify, and can therefore *inform* them *about*, the (distal) patterns in the mud on the ground. Depending on their previous experience, detecting these patterns might *inform* them *about* the nearby presence of a deer, even if they cannot see it at that moment. This way people (and other animals) can

plan and guide their behaviour in relation to an expected situation (the presence of a deer), in ways that might be influenced by past, remembered situations (previous hunting expeditions, stories told and heard about similar hunting situations, and so on) and also by imagined situations.

4.3.3 Behaviour as dynamic pattern formation

The hypothesis put forward by Haken (1996) and Kelso (1995) is that the *detection* of spatiotemporal patterns in the flowing energy arrays, i.e. perceiving the environment, is achieved by the *formation* of spatiotemporal patterns of neuronal activity. This involves moving the body parts that compose the perceptual systems around to sample the ambient arrays (J. J. Gibson, 1966, 1979/2015). These global patterns of neuronal activity are enabled by the constructive interactions that occur in the sensory surfaces, and how they become integrated in the on-going activity of the rest of the brain-body-environment system. In the case of visual perception, pattern detection/formation might result from the perturbations, i.e. the structured flux of photons arriving at the retina, to the ongoing (neuro)physiological dynamics of distributed neuronal assemblies.

Sensing a pattern available in the structured ambient array corresponds to settling on an attractor. Making this theoretical suggestion does not explain all aspects of what we might want to know about perception. Rather, it offers an alternative starting point for theoretical and empirical studies. The study of spontaneous pattern formation in nonequilibrium systems, as formalised using the mathematical tools of dynamic systems theory, offers a metaphor or framework that is an alternative to the representational-computational framework, one that is consistent with Gibson's notion of perceptual systems 'tuning in' or 'sensing' or 'resonating to' the patterned energy and material arrays available in the environment by which animals might pick up the invariants that specify relevant environmental features.

Similarly, the *production* of spatiotemporal patterns of body movements, i.e. controlling overt behaviour, is also achieved by the *formation* of spatiotemporal patterns of neuronal, and neuromuscular, activity (Kelso, 1995). This involves perceiving the relevant affordances. These global patterns of activity are enabled by the constructive

interactions that occur in the motor surfaces, and how their (feedback) effects in the sensory surfaces become integrated in the on-going activity of the rest of the system.

The cognitivist view commonly describes behaviour in term of a linear sequence such as sense-compute-act. From the radical embodiment view, this suggestion makes no sense. It is certainly the case that some specific perceptual exploration may occur *before* some specific course of action. For example, a commuter might see the train arriving at the platform before boarding it. However, visually perceiving the train depends on her moving her eyes, head, etc, to sample the visual flowfield and detect the invariants that specify the arriving train. Thus, *perception involves action*. On the other hand, successfully boarding the train depends on her perceiving the position of the train relative to her own body and on her skilfully moving different body parts relative to each other and relative to the ground and the train door. Thus, *action involves perception*. What this illustrates is that perceiving and acting are aspects of the same behavioural flow, always occurring simultaneously in sensorimotor loops rather than in linear logical, computational sequences of sense-compute-act. In the next section I focus on the relation between abilities to perceive and act and affordances.

4.3.4 Abilities and affordances are relational and dynamical features of the developmental system

The notions of abilities and affordances are intimately related. Abilities refer to what the animal can perceive and do in its environment, and therefore depend on the organization of the animal's body (the current product of its ontogeny) as much as they depend on what the current situation affords. On the other hand, affordances refer to the behavioural opportunities that the current situation provides the animal, and therefore depend on the structure of the environment (the current product of its history) as much as they depend on the animal's current abilities. Chemero (2009, p. 151) argues that "affordances and abilities are not just defined in terms of one another… but causally interact in real time and are causally dependent on one another."

This view of affordances and abilities is thoroughly relational and processual. It is sensible, for the purpose of communication, to say that 'the animal' has the ability to *X*, but this is not accurate and must be qualified. Abilities are not something that animals 'have' in the same sense that they might have four legs and a tail, for example. Rather,

abilities are here understood as relations inherent in the animal-environment system. When we say that some animal 'has the ability to *X*', what that means is that the animal has the relevant bodily organization on which the abilities depend. On the one hand this includes bones, joints, tendons, muscles, and so on, and the biomechanical properties they instantiate. On the other hand, this also includes neuromuscular networks and the coordinative properties they instantiate. Note that saying that an animal has the bodily organization that supports the ability to *X* plays a similar epistemic role to saying that it has the 'information' (meaning representations) required to solve some task, however framed from the perspective of radical embodiment.

The counterpart to abilities are affordances. Gibson and some of his followers argued that affordances are properties of the environment (J. J. Gibson, 1979/2015; Turvey, 1992; Turvey et al., 1981). More recently, Chemero argued for a dynamical account in which "affordances are relations between abilities to perceive and act and features of the environment" (Chemero, 2009, p. 150). Because an animal's abilities and the features of the environment are not static but change over time, I agree with Chemero that a dynamical and relational account is more appropriate here. The dynamics of the affordances available to animals can be studied over developmental and behavioural timescales. Additionally, the dynamics of the common affordances available to a group of individuals can also be studied over evolutionary timescale, for example in terms of the similarities and differences in the affordances that appear across genealogically related lifespans.

We say that an animal has the ability to perceive *A* (or the ability to do *B*) when, given that the environmental situation affords perceiving *A* (or doing *B*), the animal might perceive *A* (or do *B*). Note that this is not a lawful but a normative regularity. That is, even when the appropriate situation is the case, the animal might not *A*. I have the ability to climb stairs and walk through doors, but I do not climb every staircase and walk through every open door I happen to come across. More generally, at any moment animals are likely to find themselves in situations which afford several different behaviours but end up exhibiting one behaviour rather than the others. The reason for this is that there are other determinants of behaviour such as the intentions of the animal as well as random or unpredictable fluctuations.

4.3.5 The role of intentions from a dynamical perspective

I explicitly include intentions as playing relevant roles in the determination of the flow of behaviour. However, intentions are here conceived without invoking representations. Juarrero (1999) suggested that intentions could be understood in dynamic terms, as global patterns of firing activity in distributed local networks of neurons (neuronal assemblies). A prior intention (to do A rather than B, C, or D) is formed in the course of moving about, in response to what is being perceived as well as the animal's overall physiological state such as the animal is asleep or awake, hungry or satiated, cold or comfortable, and so on. Once formed, it become integrated in the rest of the neural dynamics, making some behaviours more likely to appear than others.

Many intentions can be instantiated in the same period, their corresponding neural activities cooperating and competing in the global coordination dynamics. Formulated in words, an intention can be anything from vague (e.g., 'forage') to specific ('go to that patch of fruiting tree') to even more specific ('eat this fruit now'). Even though the examples above are represented as written orders, intentions are not assumed to be algorithmic instructions but dynamic patterns of firing activity in distributed neuronal assemblies.

In a similar theoretical suggestion, Kelso and colleagues argued and presented empirical evidence that intentions play the role of specific control parameters perturbing – stabilizing or destabilizing – existing behavioural attractors i.e. preferred modes of coordination (Kelso, 1995; Scholz & Kelso, 1990; Schoner & Kelso, 1988).

4.4 What produces the flow of behaviour?

I can now turn to one of the main points of this chapter, bringing together the notions of abilities, affordances, and intentions. The flow of behaviour is here conceived of as resulting from the moment-to-moment interaction among several processes that are co-occurring as aspects of the flow of ontogeny. These processes do include the activity of the nervous system, but the nervous system alone does not produce behaviour. Specifically, I suggest that the flow of behaviour results from the interplay of:

1. The (changing) environmental situation in which the target animal is involved and what it affords given the animal's abilities. The situation might include the presence and activities of other animals, and the product of those activities.
2. The (changing) current organization of the animal's body and what abilities to perceive and act they support given the environmental situation.
3. The animal's (changing) prior intentions, which can be more or less specific. The local dynamics that correspond to prior intentions might change the prior probabilities of different behavioural attractors and therefore influence what the animal ends up doing. Alternatively, this can be framed with a stronger reference to the environment. Different aspects of the environment afford different behaviours, and they differ in terms of valence, i.e. in terms of how strongly they invite or discourage the animal to perform the corresponding behaviour.
4. Random or unpredictable fluctuations in any of the above processes.

All determinants listed above should be taken as processes, i.e. as persisting under transformation in time. Each of these constrains and is constrained by the others, and it is in this history of constructive interactions that the flow of behaviour is continually being produced. Conceived of as an aspect of the flow of ontogeny, behaviour is therefore a historical, relational, constructive, and contingent process.

One consequence of this view is that the analysis of any behavioural event should assume, as its starting point, that the animals involved already have a certain organization (the result of their ontogeny thus far) and are already actively engaging with their environment as a function of their existing abilities and intentions, as well as the affordances of the environment, and the encompassing history which has brought about the current situation under analysis.

Every behavioural event is a unique occurrence in space and time. However, different events can share enough similarity or family resemblance to justify grouping them in the same category and giving it a name, allowing for scientific generalisation. Thus, even though no two people ever walk in exactly the same way, and even though no two steps are ever exactly identical even for the same person, we are justified in speaking of 'walking' or 'lifting a leg' as behaviour categories that may recur throughout an animal's ontogeny and that may be instantiated in different ontogenies.

Behaviour categories can be defined by the regularities across the instances that compose them. According to Drummond (1981), these regularities include (1) the location or places where the behaviour typically occur; (2) the general orientation of the body in the environment; (3) the posture or changes in posture of body parts in relation to each other and in relation to relevant features of the environment; (4) other physiological adjustments which do not involve overt movement such as adjustments in alertness, body temperature, bioluminescence, electric discharges, glandular secretion, and so on; and (5) physical changes brought about in the environment as a consequence of bodily movement, such as the displacement of objects, the consumption of food, the release of bodily materials (e.g. pheromones, waste products, small body parts such as in moulting or cell renewal).

4.5 Learning is a possible developmental consequence of behaviour

Throughout ontogeny, animals continually perceive and act in their environment, in a seamless sequence of behavioural events. When we use terms such as walking, looking at a tree, reaching for a fruit, and so on, we are distinguishing some aspect of the animal's ontogeny that implies a specific spatial resolution (body parts, objects in the environment), a specific temporal resolution (e.g. seconds and minutes), and a specific epistemic interest (e.g. in terms of perception and action). At a finer spatiotemporal resolution, and with a focus on metabolism, these behavioural events would *correspond to* a sequence of physiological events – including neurophysiological events (Rose, 1981). On the one hand, behavioural events – and thus (neuro)physiological events – are enabled by the network of metabolic processes which compose them. On the other hand, the global behavioural dynamics – and thus the global (neuro)physiological dynamics – constrain (reduce the degrees of freedom of) those metabolic processes. This two-way relation between the parts and the whole they compose is not mysterious but a general characteristic of nonlinear dynamical systems (Haken, 1983; Kelso, 1995; Strogatz, 2015).

By reducing the degrees of freedom of 'lower-level' processes, such as cellular metabolism within neurons and muscle fibres, behavioural events may, over time, lead to systematic changes in the organization of the animal's body – especially in the organization of different parts of the nervous system. These changes can be more or less

persistent, more or less reversible, more or less influential in the global behavioural dynamics, more or less robust to perturbations which might come from the environment, from the activity of other neuronal networks, from other body parts such as bones, muscles and joints, or from local fluctuations. Such changes might, in turn, affect the animal's capacity to synthesize certain behaviours. When behavioural scientists associate such behavioural changes with improved performance in some target task under scrutiny, they commonly say that animals have *learned* something about the task from their experience.

Learning refers, not only to certain behavioural events within the animal's flow of ontogeny, but also to a specific consequence of those events, namely some change in the tested ability that we observers might consider to be an improvement. In other words, the verb 'to learn' refers, not simply to the performance of a task, but rather to the performance of a task *and* the occurrence of an anticipated outcome. To use the distinction put forward by philosopher Gilbert Ryle (1949), the verb 'to learn' is not a "task verb" but rather a "success verb" or an "achievement verb".

To see the difference between a task verb and a success or achievement verb, consider the example of a person *running* and *winning* a marathon. From start to finish, she is exercising several abilities such as steering away from obstacles, adjusting the pace, following the landmarks to stay on track, grabbing bottles of water and drinking from them, and so on. She is *behaving* in many ways, but we will only know if she *won* the race retrospectively, depending on her performance relative to her competitors. In this example, 'to run' is a task verb and 'to win' is a success or achievement verb. Winning a marathon implies both running *and* achieving an outcome – namely, arriving at the finish line before the competitors.

Similarly, a person or an animal might be engaging with a task and exercising different abilities, thus *behaving* in many ways. However, we will only be justified in saying that they were *learning* something in the process if there is evidence suggesting that some relevant change occurred as a result. In line with this rationale, learning is commonly conceptually defined in terms of developmental changes in the animal's body that underlie performance, and operationalized in terms of changes in some performance measure (Hogan, 2017). From a representational-computational perspective, the changes

that correspond to learning are considered in terms of representations encoded in the nervous system (Heyes, 1994; Hogan, 2017; Hoppitt & Laland, 2013) From the perspective advocated here, the changes that correspond to learning are considered in terms of the body parts involved in abilities, especially neuromuscular networks and their inherent coordination dynamics, but also any other body parts such as muscles and joints.

Here, then, is another main point of this chapter. Conceptually, I view learning as changes in the organization of the body parts involved in behaviour that are consequent on the animal's practical engagement with some task (broadly conceived), and that are reflected in improved performance in that task over repeated attempts. I believe this characterisation is consistent with the general idea underlying how behavioural scientists use the term, at the same time as it points to a clear theoretical divergence, with no mention of the formation of association between representations. "Learning", in the words of Maturana (1980, p. 45), "is not a process of accumulation of representations of the environment; it is a continuous process of transformation of behaviour through continuous change in the capacity of the nervous system to synthesize it".

Similarly, Gibson characterises perceptual learning as "the improvement of perceiving with practice and the education of attention" (J. J. Gibson, 1979/2015, p. 243). When learning occurs, the animal has become "attuned to [environmental] information of a certain sort. The [perceptual] system has become sensitized. Differences [in the ambient arrays] are noticed that were previously not noticed. Features become distinctive that were formerly vague" (id).

A note on learning and memory

Learning is intimately associated with memory, i.e. with how past events can affect present behaviour. From a representational-computational approach, memory is understood similarly to how computers store data. The animal's experience can lead to the formation of short- or long-term structures that represent some aspects of those experiences. These representations can be retrieved and manipulated, together with new representations coming from perception of the present situation. The animal can thus compute an output that is (hopefully) adequate to its current situation. Remembering is then understood in terms of the animal retrieving some representation of past experience.

Following the computer metaphor, this implies that some agent, an internal homunculus, must interpret what these representations mean.

The account above is inconsistent with a radical embodiment approach. An alternative way to understand memory and remembering is in terms of organizational changes consequent on experience. I have suggested above that behaviour is controlled by the historical and constructive interplay between abilities, intentions, and affordances (as well as unpredictable fluctuations). Abilities and intentions depend, among other things, on the biomechanical properties of body parts such as bones, muscles, joints, etc, as well as the coordinative dynamics of neuronal networks.

The biomechanical properties of different body parts and the coordinative dynamics of neuronal networks are a function of their organization, i.e. the relations among a set of material components that persist even as individual components might be incorporated, transformed, or eliminated (e.g., due to cellular turnover and metabolism). Experience – the flow of behaviour – can lead to developmental changes in the organization of different body parts because the global patterns constrain the local processes that compose them. Thus, the current biomechanical properties and coordinative dynamics of different body parts are linked to past experience by a history of causal relations and therefore *informed* by them.

This way, past experience can affect present behaviour – not because the animal formed representations about past events, but because past events played a role in determining the present organization of the animal's body. In sum, past experience is not 'represented' but 'embodied'. Remembering can then be understood in terms of the enactment of patterns of activity (such as neuronal activity) that are not only similar to past activity but causally linked to them. Because of this constraint between them, the current activity can *inform* the animal *about* its past experience.

4.6 Summary and conclusion

The motivation for this chapter was to present a view of behaviour from the perspective of radical embodiment, as an historical achievement of the system composed of the whole animal-in-its-environment. I rejected the theoretical suggestion that cognition is about obtaining and manipulating 'information' (meaning representations) about the

environment. In its place, I sided with the suggestion that cognition is about detecting relevant environmental information (*sensu* constraint between situations or what-it-is-we-can-learn-from-it) and using it to guide behaviour, including planning, remembering, and imagining (conceived of as enacted dynamical patterns, not computations with representations). I have also rejected the theoretical suggestion that behaviour is the expression of the activity of the nervous system and sided with the suggestion that behaviour is a dynamic pattern or a trajectory in phase space that involves the entire animal-environment system (including but not restricted to its nervous system).

The ambient arrays of light, sound, chemicals, surfaces, and so on, are rich in patterns or structure. Animals also play their part in creating such patterns by their presence and behaviour. Because the structure of the ambient arrays is linked to other situations, including past and future situations, by constraints, they can *inform* the perceiving animals *about* them. Skilled animals can thus sample and explore the ambient arrays with their perceptual systems and use the information they detect to guide their behaviour in relation to the immediate context but also in relation to situations that are distant in space (located elsewhere) and time (past and future). These patterns are publicly available and might be detected directly (i.e. without the mediation of mental representations).

Some of the information available in the environmental situation is especially relevant for behaviour because it specifies opportunities for action or affordances. Affordances refer to what abilities animals might exercise, given the situation in which they find themselves. Abilities refer to what animals might perceive and do, given their bodily organization and the affordances currently present. These notions are intimately related. In any situation there are possibly several abilities that the animal might exercise or, in other words, several affordances that the animal might engage. What it ends up doing depends on many things. I suggested that the flow of behaviour results from the dynamic, real-time interplay of (1) the current (changing) affordances of the environment; (2) the current (changing) abilities of the animal; (3) the (more or less specific) intentions of the animal (or, alternatively, differences in the valence associated with aspects of the environment); and (4) random or unpredictable fluctuations.

Finally, I distinguished learning from behaviour. I did so by defining learning as a possible consequence of behaviour, namely the occurrence of developmental changes in the body parts involved in abilities. This includes changes in neuromuscular networks and their inherent coordination dynamics, but also in any other relevant body part such as muscles and joints. In the next chapter I will explore how this view of behaviour leads to an alternative approach to social learning and culture.

Chapter 5 Living and learning socially

5.1 Introduction

In chapter 3 I presented a view of ontogeny as a dynamic, relational, constructive, and contingent process. In chapter 4 I proposed that the flow of behaviour is produced by the historical interplay between the abilities and intentions of the animal and the affordances of the environmental situation (as well as unpredictable fluctuations). I distinguished learning as a possible consequence of the flow of behaviour – not in terms of the acquisition of representations but in terms of organizational changes in the body parts involved in controlling behaviour.

Oftentimes, an animal's environment includes other animals. The bodies of animals reflect light, produce mechanical waves in the medium, eliminate substances, irradiate heat, exert mechanical forces, and so on. Thus, by their presence and activity, animals make environmental information available in the form of specific patterns in the ambient arrays of light, sound, chemicals, and so on. It is this information that skilled animals might detect (not 'represent') when perceiving and acting in relation to, or together with, others around them.

The goal of this chapter is to discuss some consequences that assuming this relational, historical, developmental systems view of ontogeny and behaviour might have for the study of how animals live and learn socially. I will first discuss the notion of behavioural coupling and its role in social living more generally before I focus more specifically on social learning.

5.2 Living socially

When two or more animals live together, however briefly, they bring their histories, their abilities, and their intentions, to the common situation in which they find themselves. If they can detect each other's presence, and if they can act towards each other, the behavioural flow of each might, as a consequence, influence and be influenced by what the others are doing; they become (statistically) correlated and (mechanistically or informationally) coupled, at least transiently (De Jaegher & Di Paolo, 2007; De Jaegher, Di Paolo, & Gallagher, 2010; Fogel, 1993; Kelso, 1995; King, 2004; Maturana & Varela, 1992).

Behavioural coupling can involve mechanical forces, such as when parents carry their offspring around. It can also involve informational relations by means of perceiving each other, even if animals are not in direct physical contact. Informational relations do not refer to acquiring representations but refer to the existence of (lawful or normative) regularities among situations. For example, a skilled predator might detect the patterns in the optic array that specify the presence of a prey. According to its abilities, it can use the optic form that specifies the prey's body to steer locomotion towards it, until the prey is close enough to be captured (J. J. Gibson, 1979/2015; Lee, 2009).

I suggest that behavioural coupling can be unidirectional or bidirectional. If the unfortunate prey, in the example above, is oblivious to the approaching predator, the behavioural coupling that is instantiated between them is unidirectional: the prey's behaviour influences the predator's behaviour but not vice versa. Suppose, on the other hand, that only the prey detects the optic (or auditory, chemical, and so on) invariants that specify the presence of a predator and rushes to a safe hiding place. In this case, the behavioural coupling between them is also unidirectional but in the other direction: predator affects prey, but not vice versa.

Now consider the case in which both animals detect each other and, as a result, each adjusts its behaviour according to what the other is doing, in real time. Suppose the prey rushes away from the predator and that the predator chases after it, each monitoring the other and adjusting their path of locomotion accordingly, in relation to each other and in relation to the obstacles and passages in the cluttered environment which they co-inhabit.

In this case, the behavioural coupling is bidirectional: the predator's behaviour affects the prey's behaviour *and* the prey's behaviour affects the predator's. One animal affects the other even if they do not exert mechanical forces on each other. They are coupled by a form of mutual sensitivity to the informational relations they both detect and produce. Other examples of bidirectional behavioural coupling include animals displaying in competition for access to females, verbal and nonverbal communication, mating, caregiving/caretaking, dancing, playing music, among others (King, 2004).

A common phenomenon in nonlinear dynamic systems is that mechanical or informational coupling between different systems can lead to the appearance of global spatiotemporal patterns of relation at the level of the dyad (or group). Once established, these global patterns constrain (reduce the degrees of freedom of) further local behavioural encounters among the individual participants (De Jaegher & Di Paolo, 2007; Fogel, 1993). Individuals become more likely to do some things rather than others given the wider relational and historical whole which their ontogenies help compose.

5.2.1 Consequences of living socially

There are four possible consequences of living together that follow from the above that I want to emphasise. The first consequence is rather general. Because the flow of behaviour of a target animal might be constrained by what others are doing, living together (as opposed to isolated) might affect the abilities to perceive and act the target animal ends up developing. This underlies the notion that learning can be a 'social' rather than an 'individual' process.

The second consequence is slightly more specific. By perceiving the behaviour of others and the effects produced in the world, the target animal might be *informed about* the affordances being (intentionally or inadvertently) demonstrated, which they might otherwise have missed or taken longer to detect. The target animal might thus become more likely to engage in similar tasks to the one observed. Provided their bodies are similar enough (for example, because they are closely related by genealogy and undergo similar ontogenetic trajectories), the target animal might develop similar task-related abilities to those of the observed animal. In this scenario, the development of similar abilities depended on how the target animal was influenced by the presence and behaviour

of the other animal. This is commonly considered as a case of "social transmission" or "cultural transmission" of behaviour (Hoppitt & Laland, 2013).

The third consequence I want to point out is related to the problematic notion of culture. It is also related to learning, but it involves a longer time scale than that implicit in the second consequence I just mentioned. Animals living socially can continually shape or co-regulate each other's behaviour by perceiving and responding to what others are doing in their common environment. Such recurring behavioural feedback loops can constrain developmental changes such that animals living together might become sensitive to their history of interactions. They may acquire habits (behavioural patterns described with reference to the individuals) that lead to the appearance of normative patterns of social relations within the dyad or group (Fogel, 1993; Lave & Wenger, 1991; Maturana & Varela, 1992).

For example, when newcomers join an established research group for their graduate studies, their individual behaviour influences, and is influenced by, the behaviour of their colleagues. By participating in activities such as collecting and analysing data, going to journal clubs, presenting at conferences, and so on, students are gradually introduced to the normative ways of doing science. If all goes well, they become skilled at the relevant practices that characterize the community of researchers in which they participate. In time, they move from the peripherical position of "newcomers" to the more central positions of "old-timers" (Lave & Wenger, 1991).

When a mutual sensitivity is established among partners, their individual ontogenies no longer flow independently from each other. Rather they become systematically correlated in "histories of mutual engagement", as Oyama (1999) put it. Similarly, Maturana and Varela (1992, p. 193) argue that animals living together in this way instantiate "a particular internal phenomenology; namely, one in which the individual ontogenies of all participating organisms occur fundamentally as part of the network of co-ontogenies that they bring about". I argue there is no a priori reason why we should assume this phenomenon is restricted to relations among conspecifics. After all, organisms typically live in multispecies communities and might influence, and be influenced by, each other to varying degrees. The fact that we scientists might classify them in different taxonomic groups does not impede them from perceiving and interacting

with each other as they go about living. Thus, living socially can imply relations among different species. Other examples include patterns of relation that arise between playmates, friends, mating partners, lovers, predators and prey, parents and offspring, co-workers, employers and employees, dogs and dog owners, shepherds and their animals.

Different groups of animals (including people) vary in the characteristics of their individual flows of behaviour, in the contingencies of their historical interaction, and in the global dynamics they create. Different normative patterns of behaviour might appear in different groups as a result. It follows that animals that participate in the same group might show stronger behavioural similarity compared to animals in different groups. I use the term 'groups' loosely here, as a heuristic term that indicates, on the one hand, the existence of historical relations among their components and, on the other hand, the focus of interest of the researcher. This is one aspect of the intricate notion of 'culture' (Descola & Pálsson, 1996; Geertz, 1973; Ingold, 2000, 2011a; Kroeber & Kluckhohn, 1952; Latour, 1993).

5.3 Communication as co-regulation of behaviour

Communication is a case of special interest in terms of how animals make environmental information publicly available to be perceived by others with already tuned perceptual systems. For some purposes, such as engineering, it might be useful to treat communication as the systematic correlation of signals between a source and a destination regardless of what they might mean (Shannon, 1948a). However, for the purpose of studying behaviour, this view of communication is arguably less useful analytically. Similarly, but with an explicit interest in what signals might mean, verbal communication in humans is commonly characterized in terms of the transmission of 'information' (meaning 'representation' or 'knowledge') (e.g., Csibra & Gergely, 2009).

However, to be consistent with a relational, historical perspective, here I side with a view of communication as a continuous process of co-regulation of inherently meaningful behaviour (Fogel, 1993; King, 2004; Thompson, 2007). Behaviour is inherently meaningful in the sense that it is associated by (lawful or normative) constraints with the situations they involve, including past and future situations, as well as the broader context in which they occur.

This notion of communication applies to any form of coordinated behaviour, not only human-like language. The ability to use articulate language, such as seen in humans, expands the possibilities of communication but does not inaugurate it. Communicative abilities persist under transformations from the early stages of the lifespan until death, as the bodily structures that enable them persist and change. Vocalising, gesturing, speaking, writing, drawing – all these behaviours structure the ambient arrays of light and/or sound in meaningful ways and can play central roles in how humans and other animals behave and acquire skills.

When animals vocalize, their actions produce mechanical waves in the medium (air or water), thus creating environmental information that is publicly available. Because such sound patterns are lawfully associated with the actions that produced them, they can *inform* skilled listeners *about* the presence of the vocalizing animals. Animals with appropriately tuned auditory systems and that happen to be close enough might detect the sound patterns. Depending on their own abilities and intentions, listeners might use this environmental information to, for example, adjust their path of locomotion towards the sound source. This way, predators can find prey and conspecifics can stay within close proximity as they forage.

Suppose that animals commonly vocalize upon finding a food source, as is the case in many species of primates (Rapaport & Brown, 2008). In this case, the sound patterns that appear in the environment are not only lawfully associated with the actions that produced them but are also normatively associated with the presence of the food source that motivated those actions. Thus, the sound patterns can *inform* skilled listeners *about* the presence of the vocalizing animal and also about the presence of food. It might be that, initially, the young animal is motivated to move in the direction of the calls with the intention to be reunited with its mother. However, by moving to the new place, it might also learn that the objects it sees her manipulating there afford eating. Additionally, in such situations the young will also be exposed to the normative association between the specific vocalization and the presence of food. Thus, if the animal is able to abstract (detect higher-order relations among different situations), it might discover that those sound patterns *mean* or indicate food. Later on, in similar situations, this knowledge or awareness can play a role in guiding their behaviour.

Similarly, the movements in the vocal tracts of human speakers and singers produce mechanical waves in the medium, the movements of the hands of human writers and drawers produce traces of pigment on a surface, and the movements of the fingers of a typist produce patterns of contrasting pixels on the monitor. It is the environmental information thus made available in the ambient arrays of sound and light that might be picked up by skilled listeners, viewers, and readers (J. J. Gibson, 1966). The meaning of such patterns of sound and light, just as in the case of primate food calls, cannot be transmitted. Rather, they must be discovered. This depends on, at the same time as it sustains and potentially transforms, the abilities developed by the animals within their history of engagements in similar communicative contexts.

Imagine two adults are fluent speakers of the same language and one intends to communicate a 'piece of knowledge' to the other by saying that '*X* is the case'. The speaker's behaviour creates mechanical waves in the environment that the listener might detect. These waves are linked to the speaker's movements by lawful constraints and therefore can *inform* the listener *about* those movements. Moreover, the movements of the speaker are linked to her communicative intentions by a set of causal processes and by normative constraints (the norms of language use). Therefore the sound waves can potentially *inform* the listener *about* the speaker's communicative intentions, too (Dretske, 1981). However, to interpret the meaning associated with the sound waves, the listener would have to be aware of the constraints linking the speaker's intentions and the form of the communicative behaviour. This is not straightforward. Detecting the sound patterns is not the same as making sense of what they mean. The listener has no alternative than to hear and make sense of the sounds she hears *according to her own abilities*, which depend on the contingencies of her own prior history in similar situations. Thus, the meaning of what she hears might be quite different to the meaning of what he spoke and intended her to hear.

However, speakers and listeners, or writers and readers, are not isolated individuals but part of the same group (say, British nationals), or possibly part of different groups correlated by historical contingencies (say, British and US nationals). Therefore, the communicative abilities people already have when they eventually meet – their ways of speaking or writing – might be functionally similar. These abilities need not be identical

but need only be similar enough to create a common ground, such that the sound or light patterns might have roughly equivalent meanings to each partner. This common ground allows the communication event to unfold in relative harmony. In this case, the individual flow of behaviour of those involved in the communicative situation might seem more or less coherent (to themselves as well as to us observers). In such situations, we might feel we understand what the other is saying – at least enough to keep the conversation going.

5.4 An ecological view of learning socially

In his last book, James Gibson writes about how living socially might influence learning. The following passage (J. J. Gibson, 1979/2015, p. 246) is worth quoting in full:

> The child becomes aware of the world by looking around and looking at, by listening, feeling, smelling, and tasting, but then she begins to be *made* aware of the world as well. She is shown things, and told things, and given models and pictures of things, and then instruments and tools and books, and finally rules and short cuts for finding out more things. Toys, pictures, and words are aids to perceiving, provided by parents and teachers. They transmit to the next generation the tricks of the human trade. The labors of the first perceivers are spared their descendants. The extracting and abstracting of the invariants that specify the environment are made vastly easier with these aids to comprehension. But they are not in themselves knowledge, as we are tempted to think. All they can do is facilitate knowing by the young.

Writing in the 1970s, it is probably not surprising that Gibson referred to humans only. However, his ideas, as well as the others reviewed above, suggest the basis for an ecological approach to social learning that might be expanded to other species. In this passage, Gibson acknowledges that learning can be influenced or facilitated by other persons such as parents and teachers who, by their activity, make it easier for children to extract (detect) and abstract (detect higher-order relations among) the invariants that specify features of their common environment. While it makes sense to describe this, in loose terms, in terms of *transmitting* "the tricks of the human trade", he goes on to note that knowledge is *not* transmitted. Rather, parents and teachers act as facilitators in the processes that might lead the young to learn about the world in particular ways. Two paragraphs later he seems to contradict himself by referring to "culturally transmitted knowledge" (p. 247). Later, when writing about how language mediates knowing, he says

that “We transmit information and convey knowledge. Wisdom is handed down.” (p. 248).

The choice of words might be somewhat inconsistent. However, Gibson characterises this transgenerational phenomenon in terms of “ways to facilitate knowing” without recourse to representations. His use of ‘transmission of information’, in this context, is related to verbal and written language. This probably reflects the common-sense use of the term information and the mathematical theory of communication (Shannon, 1948a). “But we should never forget”, Gibson adds, “that this is information that has been put into words. It is not the limitless information available in a flowing stimulus array” (J. J. Gibson, 1979/2015, pp. 248-249).

Words are linked to the invariants they indicate by the conventional regularities established among language users. Thus, words can *inform* skilled speakers-listeners, whose linguistic abilities have been shaped by those very regularities, *about* those invariants. Thus, by perceiving the words and pictures that others produced, we might become aware of the features of the environment they refer to. As Gibson argues, words and pictures allow people to perceive, and gain knowledge ‘at second hand’, as contrasted with perceiving and knowing ‘at first hand’. The latter requires that the perceiver-knower extracts the relevant invariants from the ambient arrays by him- or herself. The former implies that the relevant invariants have already been extracted and put into words or pictures.

In the previous chapters and in the first part of the present one, I have laid out some of the main ideas that form the basis for a processual, developmental systems view of organisms, and a radical embodiment view of cognition. So far, I have said little about the current literature on social learning and it is finally time to do so. In the second part of this chapter, I elaborate on how the insights developed thus far might be used to study social learning in humans and nonhuman animals from an alternative theoretical framework. Then, in Chapters 6 and 7 I will report two empirical studies I conducted motivated by this framework.

5.5 Categories of social learning

Drawing on an earlier, influential definition by Heyes (1994), Hoppitt and Laland (2013, p. 4) defined social learning as "learning that is facilitated by observation of, or interaction with, another individual (or its products)." Social learning comes in many flavours and different typologies have been proposed – see, for example, Galef (1988); Heyes (1994); Hoppitt and Laland (2013); Whiten and Ham (1992); Whiten et al. (2004). These works are extremely valuable in that they help synthesize and make sense of an enormous literature spanning over a century.

Table 5-1 provides the definitions of twelve categories of social learning processes endorsed or proposed by Hoppitt and Laland (2013, all definitions quoted below were taken from p. 64). I also included the suggested mechanism thought to bring about learning in each case, as well as the typical behavioural events they involve, described in simple (theoretically poor) terms. Teaching is absent from this table as the authors argue it is orthogonal to learning (I will return to this below).

The first thing to note is that some categories provide rather generic descriptions/explanations while others are more specific. For example, 'social facilitation' is defined in broad terms as occurring when "the mere presence of a demonstrator affects the observer's behaviour". In contrast, 'observational conditioning' is defined more specifically as "a subset of stimulus-stimulus learning in which observation of a demonstrator exposes the observer to a relationship between stimuli at t_1, and exposure to this relationship effects a change in the observer detected, in any behaviour, at t_2."

The second thing to note is that some definitions are clearly tied to a specific theoretical framework, namely the associative learning theory, where in others this is not the case. This is reflected in the different mechanisms proposed as explanations. The third thing to note is that, often, more than one 'mechanism' is proposed as an explanation for the same category of social learning process. The fourth point to note is that the categories are not always mutually exclusive, such that the same observed phenomenon might fall within more than one category.

The authors themselves discuss some of the points I raised above and note that their primary intention was to clarify how the terms are being used in the literature rather than to provide a pristine classification scheme. Terms like imitation, emulation, and stimulus enhancement acquire a more or less precise meaning from the specific experimental or observational context in which they are invoked as descriptive or explanatory categories. Some imprecision is to be expected among authors and papers and can even be constructive in the continuous constitution of scientific knowledge. Ambiguity and misunderstandings might be avoided by pointing to the specific epistemic practices which provide the context and frame of reference for interpretation in any case.

Thus, by raising these issues I do not mean to be criticizing their efforts. Rather, I mean to indicate that typologies such as this reflect the history of the discipline and are thus written from a theoretical perspective that is not fully consistent with the one I am developing in this thesis, especially due to their commitment to the theoretical notion of representations and its role in controlling behaviour. For example, one possible mechanism of 'response facilitation' is the "priming of brain records corresponding to an action" (Hoppitt & Laland, 2013, p. 69). In other cases, the mechanism is unclear or frustratingly vague. For example, in "production imitation", which is historically one of the most prominent categories of social learning, the novel action, or the novel sequence of actions, is simply said to be "acquired by the observer directly through observation"(p. 73).

Table 5-1. Classification of social learning mechanisms according to Hoppitt and Laland (2013), with definitions taken from page 64. L: learner; F: facilitator.

Social learning process	Definition	Suggested mechanism	Typical behavioural events associated with learning
Stimulus enhancement	*Stimulus enhancement* occurs when observation of a demonstrator (or its products) exposes the observer to a single stimulus at time t_1, and single stimulus exposures effects a change in the observer detected, in any behaviour, at t_2.	(1) F's behaviour draws L's attention to the situation; (2) F's behaviour makes the stimulus more accessible to L.	F behaves in relation to some situation. L's attention is drawn to the situation.
Local enhancement	*Local enhancement* occurs when, after or during a demonstrator's presence, or interaction with objects at a particular location, an observer is more likely to visit or interact with objects at that location.	(1) Stimulus enhancement of a specific location; (2) aggregation effect; (3) L tends to move around with other animals; (4) the products of F's behaviour draw L's attention to the place.	(1) F behaves in some place. L's attention is drawn to the place. (2) L is part of a group. The group moves around in the environment. L's individual movement is constrained by the group-level pattern of movement.
Observational conditioning	*Observational conditioning* is a subset of stimulus-stimulus learning in which observation of a demonstrator exposes the observer to a relationship between stimuli at t_1, and exposure to this relationship effects a change in the observer detected, in any behaviour, at t_2.	A variation of Pavlovian conditioning.	F behaves. L observes F.
Response facilitation	*Response facilitation* occurs if the presence of a demonstrator animal performing an act (often resulting in reward) increases the probability of an animal that sees it doing the same.	The "priming of brain records corresponding to an action" (p. 69).	F behaves. L observes F.
Social facilitation	*Social facilitation* occurs when the mere presence of a demonstrator affects the observer's behaviour.	F's presence reduces L's neophobia.	F and L are present at some place. L behaves differently than it would otherwise.
Contextual imitation	*Contextual imitation* occurs when, directly through observing a demonstrator perform an action in a specific context, an observer becomes more likely to perform that action [which might or might not be novel for the potential learner] in the same context.	A variation of stimulus-response learning.	F behaves. L observes F.

Table 5-1 (continued).

Social learning process	Definition	Suggested mechanism	Typical behavioural events associated with learning
Production imitation	*Production imitation* occurs when, after observing a demonstrator perform a novel action, or novel sequence or combination of action, none of which are in its own repertoire, an observer then becomes more likely to perform that same action or sequence of actions.	Unclear. The novel action, or the novel sequence of actions, is said to be "acquired by the observer directly through observation" (p. 73).	F behaves. L observes F.
Observational R-S learning	*Observational R-S learning* is a subset of response-reinforcer learning (R-S) in which observation of a demonstrator exposes the observer to a relationship between a response and a reinforcer at t_1, and exposure to this relationship effects a change in the observer detected, in any behaviour, at t_2.	A subset of response-reinforcer learning	F behaves. L observes F.
Emulation	*Emulation* occurs when, after observing a demonstrator interacting with objects in its environment, an observer becomes more likely to perform any actions that have a similar effect on those objects.	(1) L "could try and recreate the movements of objects with which the demonstrator interacted; this is termed object movement re-enactment" (p. 77); (2) L "could try and recreate the final state resulting from a demonstrator's behavior" (p 77-78)	F behaves. L observes F.
Opportunity providing	*Opportunity providing* occurs when the products of the behaviour of the demonstrator provide the observer with an opportunity to engage in operant learning that would otherwise be unlikely to arise - for example by providing an easier, less dangerous or more accessible version of the task.	(1) "operant learning (i.e. learning the connection between aspects of their behavior and reward)" (p. 78)	F behaves in some way that modifies task-relevant features of L's environment. L engages with the modified version of the task.
Inadvertent coaching	*Inadvertent coaching* occurs when the response of a demonstrator to the behaviour of the observer inadvertently acts to encourage or discourage that behaviour.	F's behaviour punishes or reinforces L's behaviour.	L engages with a task. F monitors L's behaviour and responds to specific behaviour events with either a positive or a negative emotional overtone.

The last point I raise is related to the behavioural events typically associated with the different categories of social learning processes. I included these data in **Table 5-1** to indicate how the different theoretical suggestions (of definitions and proposed mechanisms) relate to empirical observations described in common-sense terms implying less theory. In many cases, the behaviours associated with learning can be re-described in simple terms as 'the facilitator behaves in some way and the learner observes the facilitator'. In other cases, the literature suggests a more elaborate sequence of behavioural events that lead to learning. For example, Hoppitt and Laland (2013) identify two sequences associated with local enhancement. In the first, the facilitator behaves in some place and the learner's attention is drawn to that place. Alternatively, the learner is part of a group; the group moves around in the environment; the group-level pattern of movement constrains the individual-level pattern of movement of the learner.

The point in question here is simply that the passage from the (less controversial) empirical description of observed behaviours to the (possibly more controversial) proposal of definitions and mechanisms inevitably implies working within a theoretical framework. My intention is not simply to find new, eco-friendly, terminology to refer to the same categories listed in **Table 5-1**, but to revise how the phenomena of interest can be understood. What I want to do is to explore what difference it might make to start from an alternative perspective that is motivated by, and hopefully consistent with, a processual, developmental systems perspective of living systems, and a radical embodiment perspective of cognition. This is the goal of the remainder of the chapter.

5.6 Characterizing cases of social learning: an initial suggestion

It might be theoretically useful, as Hoppitt and Laland (2013, p. 63) suggest, "to devise a new classification scheme from scratch" and avoid some of the issues listed above. Arguably, a theoretically useful classificatory scheme should use a set of clearly defined criteria and use them to identify similarities and differences among cases of social learning grouped in different classes. While it is too early to propose a new typology, I would like to take a first step in this direction. I suggest some aspects of the phenomena that fall under the general category of social learning that might inform a future classification scheme. Then, I turn to teaching in the last part of this chapter.

A quick note on terminology. I will refer to the target animals or persons, i.e. those whose performance in a task of interest is being tracked, as the 'potential learners' (or 'learners', for brevity). This captures the idea that learning is a possible consequence of behaviour more clearly than the more common term observer. The term observer is misleading because it refers to the performance of a behaviour – when learning is not a behavioural category but a developmental consequence of behaviour. It is also misleading because learning can be influenced in other ways than by perceiving another animal perform a task, for example, by engaging in joint activity or by staying together and/or moving together in a common environment. I will refer to the animals or persons influencing the potential learners as the 'facilitators' to indicate their overarching role in the process. The more common term demonstrator indicates one out of several possible cases of roles that facilitators might play, many of which do not involve performing the target task. While mainly a matter of personal taste, I believe these terms capture the spirit of common definitions of social learning with more precision.

5.6.1 Clarifying the scope of the study

I suggest three criteria which can be used to clarify the focus of the study of social learning: the time scale, the spatial scale, and the outcome of interest. I expand on each below.

Different processes can have different rates or duration. Therefore, the choice of the time scale affects the number and type of events included in the study with possible consequences for how we theorize about them. For example, consider the study about nutcracking in capuchin monkeys, reported by Fragaszy et al. (2017). When a skilled capuchin monkey engages in cracking palm nuts by striking them with a stone hammer, younger monkeys that happen to around the facilitator become more likely to spend time near an anvil, more likely to manipulate nuts (and less likely to manipulate other objects), and more likely to percuss a nut on a hard surface and to strike a nut with a stone. All these influences on the behavioural flow of the potential learners might have consequences related to their learning the target task.

The units abstracted from the behavioural flows included in this study lasted a few minutes and included only a few behavioural events. Similarly, many lab-based studies

also involve analysing short periods of ontogeny (such as in this thesis, chapters 6 and 7). On the other hand, researchers might be interested in much longer time scales extending over years and involving many more relevant behavioural events. This is illustrated by developmental studies in primate tool use (Corp & Byrne, 2002; De Resende, Ottoni, & Fragaszy, 2008), anthropological studies of learning in children (Toren, 2008, 2012) and studies of learning as resulting from participation in communities of practice (Lave & Wenger, 1991).

Different processes involve different components. Therefore, the choice of the spatial scale also affects the number and type of events included in the study with possible theoretical consequences. Apart from the potential learner, the study might focus on only a few task-relevant objects in the environment. Alternatively, it might include other animals playing the role of facilitators. For example, young capuchin monkeys might encounter places which include the products of past nut-cracking activities, such as pieces of broken nuts and rocks that were used as anvil or hammer. Potential learners can explore those places according to their abilities to perceive and act and, in the process, they might learn what stones and broken nuts afford, even if no animal is performing task-related actions in that period. Alternatively, the researcher might decide to include more places, objects, and other animals as relevant components in the developmental system of the potential learner.

Learning is a possible (developmental) consequence of the behavioural flow. Therefore, researchers must define what is the outcome of interest. What is the target task (broadly conceived) that researchers expect to be learned or identify as being learned? For example, research might focus on *perceptual learning*, i.e. on how animals change with respect to their abilities to detect environmental information (E. J. Gibson & Pick, 2000; Kellman, 2002). Alternatively, it might focus on *skill learning*, i.e. in how animals change with respect to their abilities to engage in a target task such as sawing a plank (Ingold, 2011b) or using a stone tool to crack open hard fruits (Resende, Nagy-Reis, Lacerda, Pagnotta, & Savalli, 2014). In *place learning*, animals explore the affordances of different places and, as a result, become oriented to, and able to navigate in, their environment (J. J. Gibson, 1979/2015). Another study might focus on *language learning*

(Tomasello, 2003). Clarifying the spatial scale, temporal scale, and the target task might help identify phenomena that should be grouped as cases of the same category.

5.6.2 Clarifying the 'social influence' under focus

Social learning is a possible developmental consequence of the influence that facilitator (or facilitators) might exert on the learner. Hoppitt and Laland (2013, p. 5) suggest that "a researcher's primary task is to investigate the manner in which learning is social, rather than to answer whether learning is social or not". I suggest five criteria to characterize social influences on learning. These criteria refer to (1) the environmental information the facilitator makes available by means of its presence and behaviour; (2) the general role played by the facilitator in the learning process; (3) whether the influence is unidirectional or bidirectional; (4) whether the influence is direct or indirect; and (5) the prior intentions involved in the process. I expand on each below.

The facilitator's presence and behaviour create environmental information available to learners

The body of the facilitator reflects (and, in some cases, produces) light, creates mechanical waves, eliminates chemicals, and so on. These patterns in the ambient arrays are linked to the facilitator's body by regularities and can therefore inform skilled observers about the presence and behaviour of the facilitator, the affordances being engaged, and the consequences of that behaviour. If the potential learners are near the facilitators (or 'virtually' near, such as by means of a video monitor), they have the opportunity to directly perceive and use this information to guide their behaviour. However, the presence and activity of the facilitators might have persisting effects in the environment. In this case, the information thus created might be detected by the potential learners, and influence their behaviour, at a later moment even if they never meet.

In addition to investigating what environmental information is made available by the facilitator, we might also ask what information is, in fact, perceived by potential learners, and how they might use it to control their behavioural flow. We might also investigate how this influence facilitates learning the target task.

In Chapter 7, I will investigate the effects on learning of three types of environmental information created by the facilitator: (1) the optic information specifying

his manipulative actions; (2) the optic information specifying his face; and (3) the auditory information produced by his speech.

The two basic roles played by the facilitator

I suggest a pragmatic distinction between two basic roles that the facilitator might play. The first is to engage in the target task, i.e. the same task in which we are testing or probing the potential learners. For example, birds might perceive the song of conspecifics and this might influence the development of their own abilities to sing, such that they end up singing similarly (Janik & Slater, 2000). Naïve bumblebees might be encouraged to observe other bumblebees which had been previously trained to reach an artificial food source by pulling a string (Alem et al., 2016). In the study reported in Chapter 6, a human facilitator shows children how to use a novel tool in the context of using simple materials to make a toy vehicle. In Chapter 7, adult learners are exposed to videos of a human facilitator demonstrating how to solve different puzzles.

The second basic role of the facilitator is to engage in some activity other than the target task. Examples include cats bringing live prey to their offspring, thus influencing the development of their abilities to hunt (Caro, 1980), adult meerkats bringing scorpions to their offspring with different degrees of previous manipulation depending on their age (Thornton & McAuliffe, 2006), humans telling stories or providing verbal or written instruction about a task, and scientists giving talks describing the results of their studies.

The influence between facilitator and learner can be unidirectional or bidirectional

In terms of the direction of influence that is relevant for the ability under study, I also suggest two basic cases. The influence is unidirectional if the facilitator influences the behavioural flow of the potential learner, but the behaviour of the potential learner does not influence the behaviour of the facilitator (relative to the context of interest). This is implied in all categories listed in **Table 5-1**.

On the other hand, the influence is bidirectional in case the facilitators not only influence but are also influenced by the presence and behaviour of the potential learners, in which case they co-regulate each other's behaviour. Consider the case in which animals in a group tend to move around together in a common environment. Animals might explore the environment while, at the same time, attending to where the others are and

adjusting their own locomotion in response. As a consequence of this co-regulation, even if the companions do not engage in some 'target task' themselves, they might influence the places that the potential learners end up visiting, where they might have the opportunity to detect and use the environmental information available in those places and possibly improve in the task of interest.

This is observed in schooling in fish (Atton, Hoppitt, Webster, Galef, & Laland, 2012), agents in computer simulation (Van der Post & Hogeweg, 2008, 2006), and friends visiting a new town together and casually discovering a coffee shop. Incidentally, these cases also illustrate the facilitator engaging in some activity other than the target task. I will also discuss learning-type behaviour, which involve bidirectional influence, below.

The influence between facilitator and learner can be direct or indirect

The influence between potential learners and facilitator is direct if it occurs in real time through informational coupling. This is the case whenever the learner perceives (i.e. detects the environmental information specifying) the facilitator's behaviour as it occurs, which is the general rule (see **Table 5-1**). Alternatively, the influence between potential learners and facilitator is indirect if it occurs by means of the persisting effects in the environment brought about by their behaviour, which might be considered cases of niche construction (Laland et al., 2001; Odling-Smee et al., 2003). For example, as they forage, primates might leave behind partially processed food items and tools which others might later perceive and explore (Fragaszy et al., 2013).

Consider how buildings can influence the behaviour of those that use them. For example, much of formal learning, in so-called industrialised or developed countries, occurs within constructed places (schools) in which potential learners (schoolchildren) meet the facilitators (teachers, the administrative staff, the cleaning staff, and so on), generating normative patterns of social relations. These places are usually designed, built and furnished with the explicit intention to facilitate learning. The physical structure of museums, workshops, assembly rooms, and maker spaces also constrain the behaviour of those who visit them, thus possibly influencing what they learn in there (Lave & Wenger, 1991; Sheridan et al., 2014).

Once constructed, these places become persisting features of the environment, and can be used by many generations. Commonly the people who visit and learn in these places will never meet with those involved in their construction, and most of them will never learn how to construct buildings themselves. However, the ontogeny of the potential learners is indeed influenced by the people involved in constructing these places, as well as by those involved in keeping them in order and those whose job title explicitly includes 'teaching'. Each of these social influences play different roles and can be included or excluded in the analysis depending on the focus of interest.

The role of intentions in social learning

The question whether intentions constrain the flow of behaviour of humans and other animals, and can therefore influence learning, merits investigation. To do this consistently with the current approach, however, I suggest we must adopt a radically embodied view of intentional behaviour, i.e. one that does not involve assuming representations or a "complex mental calculator" as Caro and Hauser (1992, p. 169) put it. While a fully satisfactory account is still ahead of us, here I follow the suggestions by Kelso (1995, 2016), and especially the innovative proposal by Juarrero (1999), in which the study of complex systems play the role of a theory-constitutive metaphor.

The formation of a prior intention can be conceived of as the appearance of a spatiotemporal pattern of firing activity in a distributed neuronal assembly which alters the landscape of attractors and repellers in the behavioural phase space. This effect in the global dynamics of the animal-environment system constrains (reduces the degrees of freedom of) the possible behavioural trajectories that the system might exhibit. The appearance of an intention produces a 'contrast set' distinguishing between the behaviours that are more likely to appear (the current attractors in the behavioural phase space) and those that are less likely to appear (the current repellers). For example, when I intend to buy groceries, this impacts the global dynamics and produces a contrast set. As I walk along, I am more likely to adjust my path of locomotion and enter the supermarket or the local store (if I happen to see it first) instead of the dentist surgery or the church.

Future work should focus on proposing operational criteria to identify intentions constraining the behaviour of facilitators and learners. These criteria should allow us to

answer the following questions: (1) *What do the facilitators intend when performing the behaviours relevant to the learning process?* and (2) *What do the potential learners intend when performing the behaviours relevant to the learning process?*

In the last part of this chapter, I draw on the ideas developed thus far and focus on teaching. I will present a view that has become influential among behavioural biologists and discuss what difference it would make to investigate this phenomenon from the alternative perspective being developed. In doing so I will suggest a way to locate teaching within the context of social learning that is fundamentally different from the currently dominant view.

5.7 Teaching in humans and nonhuman animals

In their bold and influential work, Caro and Hauser (1992) provided a working definition "that attempts to unify both functional and mechanistic considerations" (p. 152), with the explicit goal to foster research on nonhuman animals. In their words (Caro & Hauser, 1992, p. 153):

> An individual actor A can be said to teach if it modifies its behavior only in the presence of a naive observer, B, at some cost or at least without obtaining an immediate benefit for itself. A's behavior thereby encourages or punishes B's behavior, or provides B with experience, or sets an example for B. As a result, B acquires knowledge or learns a skill earlier in life or more rapidly or efficiently than it might otherwise do, or that it would not learn at all.

This working definition has been rightfully praised as providing a framework with which to investigate teaching-like phenomena in nonhuman animals (Franks & Richardson, 2006; Thornton & McAuliffe, 2006; Thornton & Raihani, 2008) and it has been adopted by other leaders in the field. For example, another group of influential authors endorsed Caro and Hauser's (1992) operational definition and suggest a conceptual definition to accompany it: "we view 'teaching' as a class of behaviour patterns that are adaptations for transmitting knowledge and skills to others" (Hoppitt et al., 2008, p. 488).

One central aspect of Caro and Hauser's definition is that the facilitator modifies its behaviour in the presence of the potential learner, by which they mean "that there must

be a change in behaviour from what would otherwise take place under the same conditions if the naive observer was not present." (p. 153). In meerkats, for example, evidence suggests that the behaviour of the 'teaching' adults vary systematically in relation to the acoustic invariants picked up in the sound patterns or begging calls produced by pups of different ages (Thornton & McAuliffe, 2006).

There is, however, a second relevant sense in which facilitators might modify their behaviour in the presence of another individual. This refers to when the facilitators not only influence but are influenced by the potential learner's flow of behaviour *as they interact in real time*. In this case, both partners modify their behaviour to some extent, contingent on what the other is doing. Caro and Hauser reduce the question of sensitivity to the modification of behaviour that is performed when the other is present, arguing that a focus on a more dynamic adjustment can be ignored because "there is no a priori reason" why it should be relevant to the reproduction and survival of the learner (Caro & Hauser, 1992, p. 155). I suggest it would be more productive to treat this as a matter of empirical study rather than a priori reasoning. But even if one accepts this in the context of a comparative approach, the question of sensitivity remains relevant, I suggest even central, to our understanding of the behavioural processes underlying teaching-learning phenomena. After all, there is much more about how animals live their lives than reproduction and survival, however important these may be.

Another important aspect of Caro and Hauser's (1992) influential definition is the absence of reference to intentions. The authors offer two arguments to justify this choice. The first argument is that, "if teaching facilitates skill acquisition, then it should be favored by natural selection irrespective of the extent of attribution of mental states" (p. 169). The second argument is that, while the ability to attribute mental states to others might enhance the utility of teaching, this ability need not be taken as a requisite. A distinction could be made, they suggest, between teaching which involves the facilitator attributing mental states to other animals, and teaching "that is guided by a less complex mental calculator" (Caro & Hauser, 1992, p. 169).

These arguments point to a confusion between 'intention' and 'attribution of mental states'. The intention of a person or nonhuman animal to do *X* (rather than *Y* or *Z*) refers to a specific 'state of mind' – or, in dynamical terms, a specific neuronal dynamics – that

"play a distinctive role in the etiology of actions" (P. Jacob, 2014). I suggested above, following Juarrero (1999), that this distinctive role lies in how intentions change the layout of attractors and repellers in the behavioural phase space defined by the animal-environment system. In other words, the intention to do *X* (rather than *Y* or *Z*) changes the global dynamics of the system such that some behaviours become more likely to be performed than others.

A person or nonhuman animal need not 'attribute mental states' to other animals in order to intend to do something. Some people – notably those people whose lives unfold within the specific historical context commonly called Western – do indeed assume that the behaviour of other people (and possibly nonhuman animals as well) is guided by mental states. This habit is so ubiquitous in our ('Western') experience such that it might seem not to require justification. The representational-computational framework both thrives on, and further supports, this theoretical construct. I have been doing my best to avoid it, along with the corresponding notion that behaviour is controlled by a "mental calculator", but I welcome including intentions in the study of social learning.

For researchers interested in a functional approach, the suggestion to ignore questions of sensitivity and the attribution of mental states to others might have proven liberating. However, avoiding the question of sensitivity by an a priori argument is a poor justification and we should do better. In addition, questions about intentions should not be confused with questions about 'attributing mental states'. Moreover, students of animal teaching generally view cognition and learning as involving representations and computations thereupon. With these concerns in mind, I now want to suggest an alternative strategy with which to frame teaching in relation to social learning processes.

5.7.1 Teaching-learning as a relational, historical phenomenon

I take it that the following two related ideas are not unreasonable but instead capture central features of the phenomena denoted by the term teaching. The first is the idea that human teachers commonly have the explicit intention to facilitate learning – for example, because it is in their job description. The second is the idea that human teachers can be more or less sensitive to the changing needs of their pupils as they engage in teaching-learning relations. I suggest that explicitly bringing sensitivity and intentions back into

the discussion might prove equally productive in expanding our understanding of teaching-related and learning-related behaviour in humans and nonhumans towards a unifying framework in which to study social learning.

In common parlance, when we say that someone is teaching us, we are referring, among other things, to some sequence of behavioural events that the facilitator is performing. The behavioural events that are relevant to their teaching us and to our learning from (rather, with) them are precisely those which influence, and are influenced by, our own behavioural flow. Teaching, as learning socially, implies a relation unfolding in time and within specific environmental situations.

Additionally, when we talk about teaching, we are not only referring to a set of behavioural events, but also to a specific developmental consequence that they might have for the learner. The consequences that indicate learning are changes in some of the learner's abilities in a predicted direction which we observers judge to be improvements in the 'task' of interest. The teachers might also change in the process but, given the focus of interest, we might choose to ignore this in the analysis. Thus, it should be noted that 'to teach' is also a success or achievement verb rather than a task verb. We only know if some behaviour was indeed an instance of 'teaching' (as opposed to just behaving in some way) retrospectively, by examining whether learning has occurred in the target individual.

Two steps towards a definition of teaching-learning processes

Conceptually, I view teaching-learning processes as a subset of the broader class of facilitating-learning (i.e., social learning) process. I must now suggest what distinguishes teaching-learning relations from other ways in which learning can be facilitated "by observation of, or interaction with, another individual (or its products)" (the definition of 'social learning' by Hoppitt & Laland, 2013, p. 4). I will do this in two steps. The first step is the following. In all cases of learning socially, the behaviour of the potential learner is influenced by the behaviour of the facilitator (Heyes, 1994; Hoppitt & Laland, 2013; Whiten & Ham, 1992). I suggested above that this influence can be characterized as direct (by real-time informational coupling in perception-action) or indirect (by means of persisting environmental effects). I now suggest that the category of teaching-learning processes refers to cases involving bidirectional influences. To account as teaching, facilitators must coordinate their behaviour in space and time with respect to whatever

the potential learner happens to be doing, and the potential learners must be influenced in the abilities they develop by what the facilitator does. If the facilitator demonstrated the target task in the presence of the learner *but facing the other way*, to give a simple example, the partners would not be teaching-learning much about that task.

In other words, the facilitator performs teaching-related behaviours and the learner performs learning-related behaviours, *as each influence the other's flow of behaviour in real time*. Because this mutual influence is not incidental but fundamental to the form that these behaviours end up having, they should be conceived of as mutually implicated aspects of a unitary phenomenon, as poles within the relational, historical process of teaching-learning. This is conceptually radically different to the suggestion, by Hoppitt et al. (2008), that teaching-related behaviours are "orthogonal" to (i.e., that they are fundamentally independent from) social learning processes.

This perspective opens up empirical questions about how this mutual influence is instantiated that might not even be considered in the dominant perspective. How is this mutual sensitivity and behavioural coupling achieved, sustained, and dissolved, as animals engage in unfolding teaching-learning situations? What are the mechanisms involved in the coordination of behaviour, in space and time, with respect to the contingencies of the situation which they help compose? Does behavioural coupling lead to the appearance of spatiotemporal patterns of collective behaviour which, once established, constrain (reduce the degrees of freedom of) further behavioural events at the individual level? I will examine some of these questions empirically in Chapters 6 and 7.

I now take the second step in defining teaching. I do this by bringing to the foreground precisely an aspect that is neglected in the currently dominant approach. I suggest that the intentions constraining the behavioural flow of the participants is a reasonable criterion to distinguish teaching from other cases of social learning. Furthermore, it might also be a reasonable to distinguish among cases of teaching, depending on how we answer the questions I raised above: (1) *What do the facilitators intend when performing the behaviours relevant to the learning process?* and (2) *What do the potential learners intend when performing the behaviours relevant to the learning process?*

Researchers who view teaching through the lens of adaptation use terms such as ‘active’ to characterise teaching-related behaviour, as well as terms such as ‘passive’ or ‘inadvertent’ to characterise *non*-teaching-related behaviours. Note that the latter term is a common-sense antonym for ‘intentional’, which seems inconsistent with the neglect of intentions. “Use of the term ‘inadvertent’”, say Hoppitt et al. (2008, p. 490), “signifies that the demonstrator’s behaviour is not adapted to the function of transmitting knowledge, and does not imply that teaching requires any intentionality on the part of the tutor”.

It is unclear why the authors move from a statement about adaptation to one about intentionality. This might indicate a confusion between the notions of intention and intentionality. “Intentionality is the power of minds to be about, to represent, or to stand for, things, properties and states of affairs (P. Jacob, 2014). I have already mentioned a similar confusion above when presenting Caro and Hauser’s influential paper. Intentionality and the attribution of mental states are related ideas. A more appropriate expression here might be ‘intention to teach’. If this is the case, the authors would be suggesting that teaching does not require that the facilitator has the intention to facilitate learning.

The important point, however, is that questions about the phylogenetic causes and consequences of the abilities involved in teaching and learning (which are implicit in the reference to adaptation) are conceptually distinct from questions about intending to do one thing or another and should be treated separately. On the one hand, we might ask if and how a past phylogenetic history has constrained the development of the facilitator’s behaviour currently observed. On the other hand, we might also ask if and how the current intentions of the facilitator affect the teaching-learning process.

Defining teaching

I can finally suggest the following working definition of teaching-learning relations. Teaching-learning processes are a subset of social learning processes characterised by (1) the mutual influence between facilitators and potential learners and (2) the intention of the tutor to facilitate learning. All cases of teaching involve the facilitators being more or less sensitive to the presence and activity of the potential learners and adjusting their behaviour in their presence. This mutual influence is integral to bringing about learning,

and not an incidental accompaniment to it. Moreover, all cases of teaching involve the facilitators intending to bring about learning, albeit to varying degrees and in different ways that can be further investigated in each case.

5.8 Summary and conclusion

In this chapter I presented a relational, historical view of living and learning socially. Animals living socially can affect each other's environment and each other's behaviour. I referred to this mutual influence as behavioural coupling or co-regulation. This coupling can be mechanical (by physically moving one another) and/or informational (by perceiving each other). The influence between two individuals can be unidirectional (when one affects the other but not *vice versa*) or bidirectional (when each affects the other). The influence can also be direct (by real-time behavioural coupling) or indirect (by means of persisting effects in the environment). When animals live socially, their ontogenies are systematically coupled and therefore cannot be considered as independent from each other.

The contingencies of behaviour can have developmental consequences for the organization of the animal's body, including the appearance of new abilities to perceive and act. Therefore, when animals influence each other's behaviour systematically, they might be influencing each other's development. This is captured by the broad notion that learning is 'social' rather than 'individual' when it is "facilitated by observation of, or interaction with, another individual (or its products)" (Hoppitt & Laland, 2013, p. 4). I retained the general spirit of the definition of social learning but viewed it from a processual, developmental systems approach to ontogeny and a radical embodiment approach to cognition. From this view, social learning is conceived of as a possible developmental consequence that might occur when the flow of behaviour of the potential learner is influenced by the flow of behaviour of the facilitator either directly (through real-time mechanical and/or informational coupling) or indirectly (by the persisting effects of the facilitator's behaviour in the environment). The consequence in question is a change in the organization of body parts that underlies abilities that lead to improved performance in the task of interest.

Thus, conceptually, social learning processes are here taken to be relational histories of development of abilities rather than processes of transmission of behaviour. Consistent with this, while authors writing within a representational-computational framework commonly use the vocabulary related to acquiring and using 'information' (meaning 'representation' or 'knowledge'), the current approach prefers the vocabulary related to detecting and using information (*sensu* constraint between situations and specification of the environment in ambient structured arrays), abilities, affordances, and intentions.

I briefly presented a classification of categories of social learning processes, identified some of its shortcomings, and then suggested a set of criteria that might, in the future, be used to produce a principled classification. I started by suggesting that the focus of any study might be clarified by pointing to the specific time scale, spatial scale, and learning outcome of interest. I then suggested a set of criteria that might help clarify what 'social influence' means in each case. These criteria were meant to clarify (1) the role of the environmental information made available by the facilitator; (2) the basic role played by the facilitator (to perform the target task or something else); (3) whether the influence between facilitator and learner is unidirectional or bidirectional; (4) whether the influence is direct or indirect; and (5) what intentions are involved.

I finished by focusing on teaching-learning relations. I briefly reviewed an influential definition, identified some shortcomings and inconsistencies with the present approach, and suggested an alternative definition of teaching as a relational, historical process in which mutual sensitivity and the intention to bring about learning were fundamental.

In the following chapters I present the results of two empirical studies in which I investigated cases of learning socially motivated by the perspective developed here. In the study reported in Chapter 6, children visiting a science museum in Newcastle upon Tyne (UK) were invited to participate in a making activity with a skilled facilitator. I analysed some aspects of how they coordinated their behaviour with each other during the activity and attempted to investigate its association to learning. In the study reported in Chapter 7, I recruited adults to the joint eyetracking lab at the University of Edinburgh. They watched videos in which an adult facilitator demonstrated the solution to a set of

puzzles and subsequently attempted to solve them. For the main analysis, I manipulated the stimuli such that some learners could see the facilitator's face and others could not, and some learners could listen to what the facilitator was saying and others could not. I analysed aspects of how leaners coordinated their visual attention to the stimulus on the screen and its relation to learning.

Chapter 6 Learning through making together: coordinating behaviour in a joint making task.

6.1 Introduction

In this chapter, a joint making activity provides the context for a case of social learning. My primary goal is to use this example as an opportunity to flesh out, in more concrete terms, what it means to view social learning as a relational developmental process rather than as a linear transmission of information (*sensu* represented knowledge) between individuals.

Throughout this thesis, I argue for a processual, relational view of ontogeny according to which morphological and behavioural form is continually generated in real time by the set of interacting causal processes that compose the developmental system. This view was offered as an alternative to the more common view of form as pre-existing, at least partially, in the information (*sensu* instructions) encoded in the genes and expressed in development and in the information (*sensu* represented knowledge) encoded in the brain and expressed in overt behaviour. The term developmental system denotes the interdependence of the (changing) organism under scrutiny and its (changing) environment. For many animals, the environment includes tools and materials that can be manipulated in different ways. Thus, making can be understood as a process unfolding within the flow of ontogeny, as an aspect of the changes in the developmental system that manipulative behaviour can bring about. In a making activity, the maker, the tools, and the materials affect each other over time in different ways. Making is here conceived of as being achieved, not by the maker projecting a mental representation onto materials, but rather by the maker engaging with tools and materials as enabled and constrained by his

or her abilities and intentions, as well as by the physicochemical properties of the relevant set of tools and materials. In sum, making, as ontogeny, is here seen as a form-generating process rather than a form-imposing one (Ingold, 2013).

Making a toy vehicle as a context for learning socially

The Life Science Centre in Newcastle upon Tyne (UK) is a fun place to visit, especially designed to instigate the curiosity of the visitors with a set of interactive installations and supervised activities. In addition to their main exhibition area, their staff routinely conduct activities with visiting schoolchildren. In one such activity, children use common materials, such as discarded milk bottles, plastic lids, glue, and coloured paper, to produce toy vehicles. The task in this study is a simplified version of this activity. Modifications were made in common agreement between the experimenter (myself, MP) and the facilitator (the maker-in-residence at the time) so that children would engage individually rather than in groups, and a hand drill was included as a tool which we expected would be novel to the children, thus possibly allowing us to investigate implicit learning. The facilitator agreed to conduct the activity in a way as similarly as possible to the regular making activities, thus producing an interaction that was naturalistic and rich in stimuli, similar to a short-lived, small-scale apprenticeship relationship. I will refer to the adult maker as the facilitator or by using the female pronouns, and I will refer to the children makers as children, learners, or, to avoid ambiguity, using the male pronouns. I will also refer to the facilitator and the learners collectively as the participants of the study.

The way I describe the task below is inspired by a version of Russian Activity Theory presented by the social psychologist Saadi Lahlou (2011b). The task is a continuous history of causal relations in which an initial situation is transformed into a final situation. The term ‘situation’ refers to parts of the world, i.e., to how objects and materials relate to each other in specific locations. I use the term intentionally as a reference to “situation semantics” mentioned in previous chapters (Barwise, 1983; Barwise & Perry, 1981; Barwise & Seligman, 1997; Chemero, 2009; Israel & Perry, 1990). The basic idea is that situations may be linked by lawlike and/or normative constraints. Therefore, as the activity unfolds, perceived aspects of the time-evolving situation can *inform* skilled observers *about* aspects of past, present, and future situations

to which they are linked, and affect their behaviour in ways consistent with the current goal.

For example, in the making activity examined here, perceiving an attentive child and a set of unassembled, circular lids can inform the skilled facilitator that she should instruct the child to transform them into wheels by drilling holes in them. This might increase the probability that she will reach for the hand drill rather than, say, the scissors nearby. Later, perceiving the child and the assembled vehicle can inform the facilitator that she should instruct the child to decorate it. This might increase the probability that she will point at the scissors rather than at the hand drill. In addition, makers may be able to remember previously lived sequences of linked situations and use this understanding to guide present behaviour. For example, children may remember how they used a pair of scissors to cut a specific pattern from a sheet of coloured paper at school, and this might affect how they end up decorating the vehicle in the present activity. Makers may also be able to imagine (i.e., simulate perceptual experience of) possible sequences of situations and use this to guide present behaviour. For example, when asked about the steps they would take to produce a toy vehicle (see below), children may imagine a sequence of situations, informed by their previously lived experience, and use this to guide their answer.

In the current task, the components of the initial situation include the facilitator and the child with their existing abilities, as well as the tools and materials in their initial shape and spatial arrangement. The components of the final situation include the facilitator and the child, as well as the tools and materials in their final shape and spatial arrangement. Crucially, some of the materials should have been transformed to become a toy vehicle by then. It is possible that the child's bodily organization that underlies task-relevant abilities will also have transformed in the process in a manner which we might characterize as improvement in task performance consequent on practice. Although the task is a continuous process (leading to its overall goal), for analytical purposes it is segmented into a sequence of subtasks or steps (each leading to its respective subgoal). Each subtask begins with its corresponding initial situation, unfolds as this initial state of affairs is transformed into a corresponding final situation and becomes the initial situation for the following step, and so on, until the overall goal is reached, and the task is

considered completed. The unfolding situation can also be understood as the combination of two developmental systems – one focused on the facilitator and the other focused on the current child – which partially overlap during the period in which they engage in the activity together. This partial overlap indicates that the facilitator and the child may influence each other's ontogeny as they live (and possibly learn) together.

In the making activity, participants exercise their existing abilities to perceive and act as they manipulate tools and materials and influence each other's behaviour. In this process, their abilities are at once invoked and possibly transformed. For example, children might learn about what their own bodies are capable of, the properties of different materials, and how to operate the different tools. They might also learn about how to interact with other people and behave in a public space, how scientists conduct studies, how to deal with frustration and surprise, and so on. Here I focus on two possible aspects of learning, as described below. While I acknowledge that the facilitator might also change as a result of engaging in the activity, the focus here is more explicitly on the children as potential learners. Thus, the case study offered the opportunity to focus on possible roles that behavioural coordination may play in social learning.

In this activity, the facilitator's role is to assist each one of the different children. To accomplish this, she must be attentive to what each child is doing and make sense of it; she must anticipate and/or identify his or her specific struggles and provide the appropriate, context-specific support. All of this depends on her accumulated professional experience as a facilitator in similar activities and her abilities to coordinate her attention to each child's behavioural flow, adjusting her own behaviour accordingly. Similarly, children are not just making an abstract toy, but each is making their own toy in this particular day and setting, with these tools and materials, and together with this facilitator. Children bring their existing skills and preferences to the activity and explore the materials creatively while attending to the facilitator's behaviour and responding to her instructions and interventions. Thus, in a joint making activity such as this, there is plenty of room for improvisation and coordination of behaviour and attention between participants, and the skills that children already have at the start of the activity might change consequent on their experience. It is precisely this real-time coordination between

the improvised behaviour of both the facilitator and the children, and the possible association between behavioural coordination and learning, that I want to capture.

Previous studies have examined the association between visual attention and performance in tasks involving social learning. Here I will mention only two studies to highlight an important methodological aspect of the present approach. M. Carpenter, Tomasello, and Savage-Rumbaugh (1995) reported a comparative study involving imitation tasks with three captive, mother-reared chimpanzees between 3 and 21 years of age, three captive, human-reared ("enculturated") chimpanzees between 4 and 10 years of age, and six 18-month-old children. In each trial, the potential learners first had the opportunity to observe the facilitator (experimenter) perform a target action on an object ("model period"). Subsequently, the facilitator gave the object to the learners and observed their behaviour, sometimes prompting and directing them to perform the target action ("response period"). The authors operationalized joint attention episodes as the periods of time in which participants alternated between looking at the target object and looking at the experimenter's face. They reported a positive correlation between a measure of imitative performance and the average length of joint attention episodes during the response period (the corresponding result for the model period was not reported). This correlation was observed for the participants taken together as well as within each group separately. The authors concluded that "the alternation or coordination of attention to both object and E [experimenter] was associated with higher levels of imitative learning." (M. Carpenter et al., 1995, p. 231). In their study, the expression "coordination of attention" refers to how potential learners alternated their visual attention between different targets (the object manipulated or the experimenter's face) over time, but does not capture how that relates to the visual attention of the facilitator in the same period. In the present study, on the other hand, the expression "coordination of attention" refers to how potential learners alternate their visual attention between different targets (their own manipulation, the facilitator's manipulation, their joint manipulation, or the facilitator's face) explicitly *in relation to* how the facilitator also alternates between the same set of targets in the same period.

The second study I want to mention, reported in Vivanti, Nadig, Ozonoff, and Rogers (2008), used eye-tracking to compare patterns of visual attention in 18 children

with autism and 13 children with typical development, between 8 and 15 years of age. Participants were instructed to watch a set of video recordings of the facilitator demonstrating a target action and to imitate what they saw the facilitator do after each clip. Among children with autism, but not among children with typical development, the authors reported a positive correlation between a measure of imitation precision of nonmeaningful gestures and the proportion of time children spent looking at the region where the action was performed (as opposed to time spent looking at the facilitator's face or other regions in the video). While the proportion of time may be a sensible index of visual attention, it does not capture any dynamical patterns or the time-evolving relation between the visual attention of the learner and the behaviour of the facilitator. Here I examine whether these time-evolving patterns and relations are relevant to the social learning process under examination. For clarity, below I refer to the time-evolving, relational character of joint attention as *coupled visual attention*.

How to capture behavioural coordination in real time?

The point I want to make is that it is difficult, if not impossible, to capture the dynamical or historical interplay between the learner's and the facilitator's behaviour using solely measures such as the average length of joint attention episodes or the proportion of time spent looking at different targets, because variables such as those aggregate behaviour over time. As an alternative methodological approach, here I explore nonlinear methods from recurrence analysis, which make no assumptions about the size, distribution, or stationarity of the data (Webber & Zbilut, 2005). These methods permit considering temporal variability not as noise to be averaged over but as intrinsically informative about the underlying biological processes that generate the data. Eckmann, Kamphorst, and Ruelle (1987) introduced recurrence plots (**RP**s) as a tool to visualize the trajectory of a dynamical system in its phase space by means of recurrences, i.e. repetitions of a system's states over time. The RP is useful to examine the time evolution of a system because differences in the patterns that appear in the RP indicate differences in the underlying dynamical process. Moving beyond visual inspection, Zbilut and Webber (1992) and Webber and Zbilut (1994) initially suggested how some of these patterns could be extracted for recurrence quantification analysis or RQA. The cross recurrence plot (**CRP**) is a bivariate extension of the RP used to compare the simultaneous trajectories of two systems and its patterns can also be extracted and quantified using cross recurrence

quantification analysis or **CRQA** (Marwan & Kurths, 2002; Shockley, Butwill, Zbilut, & Webber, 2002; Zbilut, Giuliani, & Webber, 1998). Marwan, Romano, Thiel, and Kurths (2007) provide an extensive review of recurrence and cross recurrence analysis and Webber and Zbilut (2005) offer a more accessible introduction.

Though still relatively new in the analytical toolkit of researchers working on social learning, a growing empirical literature has applied these methods to the study of intra- and interpersonal coordination. Shockley, Santana, and Fowler (2003) reported that participants who engaged together in a conversation task coordinated their postural sway trajectories. A subsequent study suggested that this shared postural activity was mediated by convergent speaking patterns between participants (Shockley, Baker, Richardson, & Fowler, 2007). Perhaps more relevant to my discussion is another set of studies which focused specifically on gaze coordination. D. C. Richardson and Dale (2005) recorded the speech and eye movements of participants while they looked at a set of pictures of famous cast members of a TV sitcom (either ‘‘Friends’’ or ‘‘The Simpsons’’) and spoke spontaneously about them. They then played unedited one-minute segments to a separate set of participants who were looking at the same pictures and also recorded their eye movements. Speakers and listeners synchronized their gaze direction above chance with a peak at a 2 sec lag, consistent with the speaker leading the listener’s attention; and a stronger coupling was associated with higher scores in a comprehension test applied to the listeners.

Two subsequent studies are reported in D. C. Richardson, Dale, and Kirkham (2007). The first study used the same task and stimuli as that of Richardson and Dale (2005) but allowed pairs of participants to interact in real time in a two-way conversation about the cast members of the TV shows. Results showed a similar coordination except that the peak recurrence was found at lag zero. This was interpreted as evidence that participants took turns in leading and following each other’s attention. The second study reported in D. C. Richardson et al. (2007) examined live unscripted conversation between participants about a painting by Salvador Dali. Before their conversation, pairs of participants heard either the same or different passages about the painter. Results confirmed the prediction that the partners who had listened to the same passage prior to the conversation, compared to different passages, showed greater coordination of their

eye movements. Louwerse, Dale, Bard, and Jeuniaux (2012) reported that participants engaging in an unscripted face-to-face collaborative communication task synchronized their behaviour across several modalities simultaneously, including linguistic expressions, facial expressions, manual gestures, and noncommunicative postures.

Recurrence analysis seems very well suited to investigate interpersonal coordination. Therefore, applying these tools to the study of social learning might allow new insights to be gained about how learning, here operationalized as improved performance in the tests conducted before and after the activity, can be influenced by the presence and behaviour of others. However, to my knowledge these tools have not yet been used with the explicit goal to study processes of social learning. The current study is an initial effort to explore this promising possibility in a naturalistic context.

6.1.1 The present study

This study was thoroughly exploratory, guided less by the desire to do hypothesis-testing and more by the desire to explore innovative analytical tools to investigate behavioural coordination in the context of social learning while being consistent with the theoretical perspective developed in the previous chapters. I examined the dynamics of visual attention during the task and its association with learning. Children engaged individually in a making activity guided by an adult facilitator. I used cross recurrence quantification analysis to examine dynamical features of the participants' eye movements as they interacted, and I examined the children's ability to use a novel tool and to identify the crucial steps in the activity. I investigated whether the dyads synchronized their visual attention and whether their coupled attention was constrained by the time-evolving task. I also investigated whether measures of attentional coordination were associated with learning. The facilitator interacted with the potential learners in real time. Therefore, measures of behavioural coordination should be interpreted as referring to the dyad as a distributed whole rather than to the facilitator or the child as individuals.

Predictions regarding behavioural coordination

The different subtasks should impose different constraints on the participants' behavioural flows. Therefore, I expected different subtasks to produce systematically different signatures of attentional coordination (*Prediction C1*) but I do not have more detailed predictions. Given the nature of the task and the explicit commitment of the

participants, dyads should not only coordinate indiscriminately but indeed *synchronize* their visual attention during the activity in order for the joint activity to unfold smoothly. In other words, participants should tend to be gazing at the same target at the same time (*Prediction C2*). Since the facilitator explicitly agreed to follow the same predefined script with all children, I expected different dyads would produce similar temporal patterns of coupled attention (*Prediction C3*).

Predictions regarding learning.

Given that the facilitator explicitly intended to bring about learning, I predicted that children who engaged in the making activity would improve their ability to use the novel tool, compared to children who used the tool unassisted in the control condition (*Prediction L1*), and that they would learn about the steps required to solve the task (*Prediction L2*). I also examined whether coordination during the subtask involving the target tool was associated with children's performance with the tool after the activity. I expected that higher values of coordination would be associated with better performance (*Prediction L3*).

6.2 Method

6.2.1 Site and participants

The study took place at Life Science Centre in Newcastle upon Tyne (UK). A total of 42 children between the ages of 7 and 10 years were recruited from the general public visiting the museum. Of these, 17 children (8 females) participated in the making activity and were allowed to take the built product with them after. The other 25 children (15 females) provided the control condition for the implicit knowledge test. Informed consent was obtained from the children's guardians and also from the adult maker-in-residence who acted as the facilitator. The study was approved by the University of St Andrews Teaching and Research Ethics Committee.

6.2.2 Design

Most of the analyses reported refer to children who engaged in the making activity. They were first given the implicit and explicit knowledge tests, then engaged in the task with the facilitator, then repeated the tests. Additionally, as a control for the implicit knowledge test only, a separate cohort of children were first given the implicit knowledge

tests, then returned to the exhibition area and engaged in a distraction task for 10 minutes, then returned to repeat the performance tests. During this interval, children might interact with their caregivers. Caregivers were instructed to not give any assistance should their children ask about the target tool, and no instance of such assistance was observed.

6.2.3 Experimental setup

Figure 6-1 shows the setup for the control of the implicit knowledge test. **Figure 6-2** shows the tools and materials for the making activity and **Figure 6-3** shows the setup. The making activity was conducted in a dedicated room which was visually but not acoustically isolated from the rest of the exhibition area. The room included a desk with the tools and materials and participants could move around it as required. The facilitator and the child stayed on opposite sides of the desk facing each other during most of the activity. The sessions were video-recorded using a set of digital cameras. The main camera was inside a box hidden from children's view and captured the activity from the side. Participants and the experimenter also wore subjective cameras or *subcams*. These are small cameras attached to lens-free spectacles frames that capture the person's perspective (Lahlou, 2011a), and were used to code eye-gaze and to disambiguate the side-view footage when needed. Children were told the subcams would allow the experimenter to see things from their perspective, and they showed no signs of discomfort using them. The control for the implicit knowledge test involved only the novel tool and was conducted in a different space (**Figure 6-1**) close to an existing activity at the museum which provided the distraction task. In this case, the space was only partially isolated from the rest of the exhibition area. These sessions were video-recorded using two tripod-mounted digital cameras.

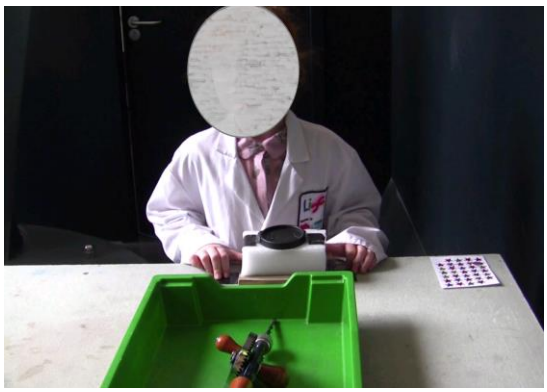


Figure 6-1. The setup for the control condition.

Figure 6-2. Tools and materials used in the making activity. For the vehicle's body, children chose among the piece of foam, milk bottle, and egg carton (left). The tools (centre) included a familiar tool (pair of scissors) and a novel tool used in the implicit knowledge tests (hand drill). The black circular plastic lids (left) were drilled to become wheels, which were then attached to the vehicle's main body using the wooden axles (left). The other materials were available for decorating: plasticine, stickers, glue, coloured paper, glitter pens, and pipe cleaners.

6.2.4 General procedure

Children were recruited and taken individually to the dedicated space. First, the experimenter (myself, MP) asked a set of questions (*Have you seen one of these before? Can you tell me what this is? What do people use this for?*) to assess familiarity with the pair of scissors (expected to be familiar) and the hand drill (expected to be novel) in counterbalanced order. The experimenter then conducted the set of pre-activity performance tests (**Test-1**). **Table 6-1** summarizes the prompts used and the responses coded. The ***implicit knowledge tests*** assessed children's ability to use the target tool. I asked them to have a go with the hand drill on a prepared target identical to what they would use during the activity. I coded whether they succeeded to drill a hole and measured the time interval from when they started operating the tool (i.e. turning the wheel handle) until success or test termination. The test was interrupted after about 30 sec of tool use without success or earlier to ensure children's safety and engagement. The ***explicit***

knowledge tests assessed children's understanding of the steps required in the task. I showed them a finished model vehicle (shown in **Figure 6-3**) and the materials available on the table and prompted them to describe the steps they might take to make a similar product. Responses were coded from the recorded videos into categories that corresponded to the subtasks defined in the script: 'pick a body', 'making the wheels', 'assembling', 'decorating', and an additional category 'other' for any other response. This resulted in an ordered categorical score ranging from zero (no subtask identified) to four (all four subtasks identified).

After the initial performance tests (Test-1), children in the experimental condition participated in the making activity, described below, which lasted for an average of 12 min (range: [9 min, 15 min]). Alternatively, children providing the control for the implicit knowledge test returned to the main exhibition area and participated in a distraction task for about 10 min. After the making activity or the distraction task, the experimenter conducted the post-activity performance tests (**Test-2**).

Table 6-1. Tests performed with the children participants before and after the activity or distraction task (control).

Test	Prompt	Responses coded
Implicit knowledge test	*Can you have a go and show me how it works?*	(1) Binary: success or no success. (2) Continuous: time to success (in sec).
Explicit knowledge test	*If you would make a vehicle like this one [Experimenter shows model], what would you have to do first? And next? And next?*	Ordered categorical: number of subtasks identified ranging from zero to four.

6.2.5 The making activity

The activity began when the experimenter finished Test-1 and the facilitator took over. The overall goal of the task consisted in making a toy vehicle under the assistance of the adult facilitator using simple everyday materials, a common tool (a pair of scissors) and a novel tool (hand drill). The task was divided into five subtasks, summarized in **Table 6-2** and illustrated in **Figure 6-3**. Note that the script is not a rigid set of rules, but the general storyline within which participants can, indeed must, improvise creatively according to the unique situation they encounter and the unique set of abilities and intentions they bring. The script *represents* the activity for communication purposes without being a *mental representation*. The script 'controls' the activity insofar as the facilitator, but not the children, committed to following it. In the first subtask (PICK_BODY), the subgoal was for the children to have chosen the material that would become the vehicle's body. The facilitator's role was to engage the children by asking them what sort of vehicle they want to make. The facilitator offered three possible materials (an egg box, a milk bottle, and a piece of foam) and prompted the children to choose one. The children's role was to communicate with the facilitator and commit to the task. This subtask ended when the child indicated his or her choice. In the second subtask (TOOL_DEMO), the facilitator demonstrated how to use the hand drill by drilling a hole in the centre of a circular plastic lid similar to what children would use to make the wheels for the toy vehicle. She also gave verbal instructions describing her actions. The

children's expected role was to attend to the demonstration. The subgoal in the third step (TOOL_USE) was to have four wheels ready to use. The children's role was to use the hand drill four times to make four wheels. The facilitator's role was to attend to the children and instruct and assist them as required for a safe and appropriate use of the tool. The subgoal in the fourth step (ASSEMBLE) was to have the basic vehicle ready. The facilitator's role was to instruct the children how to put the parts together (body + wheels + axles) and assist them if and as necessary. The children's role was to follow the instructions and assemble the parts into a basic toy vehicle. The subgoal in the fifth step (DECORATE) was the same as the overall goal of the task, namely to have the final toy vehicle ready. The facilitator's role was to show the available materials for decorating and to assist the children as needed. It was also her responsibility to keep track of time and bring the activity to an end within 5 minutes after the children started decorating. The children's expected role was to use the materials to decorate their vehicle however they wanted.

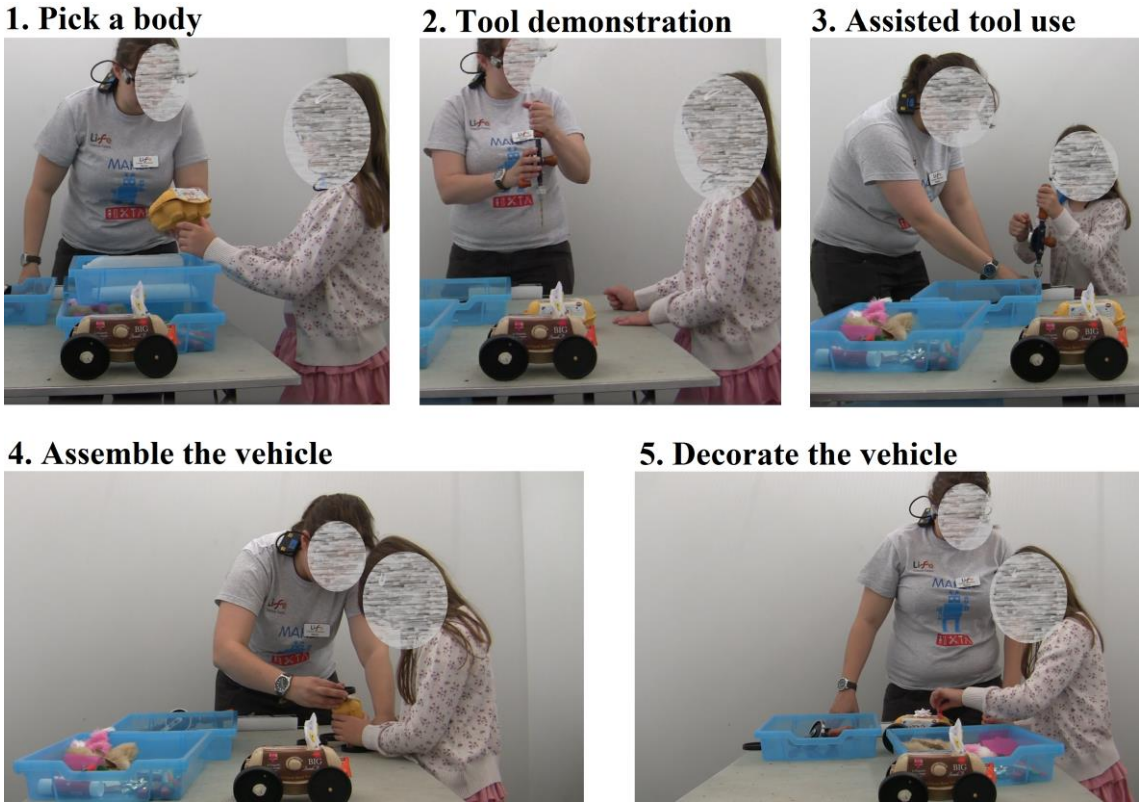


Figure 6-3. Snapshots illustrating each of the steps that compose the task. The model vehicle is on the desk. The tool used in the implicit learning test is the hand drill, shown in steps 2 and 3.

Table 6-2. The script of the making activity. Note that the final situation of one subtask becomes the initial situation of the next (shown with the same colours).

	Subtasks of the making activity				
	1. PICK BODY	**MAKE WHEELS**		**4. ASSEMBLE**	**5. DECORATE**
		2. TOOL DEMO	**3. TOOL USE**		
Initial situation	Child is recruited and ready to engage the task.	Child chooses (by gesture or verbally) one of the materials.	Demonstration lid has a hole and drill is detached from it.	Four wheels are ready, and the hand drill is detached from the fourth lid.	The basic vehicle is ready.
Changes made to the materials	Material is chosen for the body, but no changes are made to the materials.	The demonstration lid is drilled but no changes are made to the materials for the child's vehicle.	Four plastic lids are drilled and now become wheels.	Materials are put together and now have the form of a vehicle.	Other materials are added to the vehicle, depending on the children's abilities and creativity.
Final situation (subgoal)	Child chooses (by gesture or verbally) one of the materials.	Demonstration lid has a hole and drill is detached from it.	Four wheels are ready, and the hand drill is detached from the fourth lid.	The basic vehicle is ready.	The vehicle is decorated and finished.
Facilitator's expected role	Engage child. Present choice of materials for the vehicle's body. Ask child to choose one.	Demonstrate how to use the hand drill.	Prompt child to use the hand drill. Assist as required.	Prompt child to assemble the basic vehicle. Assist as required.	Offer materials and prompt child to decorate. Assist as required. Terminate activity after ca. five minutes
Examples of prompts directed to the children	*Which one of these materials do you want to use for the body bit?*	*I'm going to show you how to do it and then you're going to have a go, ok?*	*Do you want to have a go at making your wheels?*	*Can you put it together now?*	*You have five minutes to decorate your vehicle however you want.*
Children's expected role	Choose one of the materials.	Attend to the demonstration.	Use the hand drill four times to make four wheels.	Assemble the vehicle.	Decorate the vehicle.

6.2.6 Data processing

Coding and inter-observer reliability

Participants' behavioural flows were coded from the video recordings into categorical time series at a sample rate of one observation per second using the free software Solomon version beta 16.06.26 (Péter, 2016). For each session, three time series were produced for the facilitator and three time series were produced for the child. In each case, the first time series referred to their manipulative behaviour, the second to their vocalizations, and the third to the direction of their gaze. The behavioural categories are defined in **Table 6-3**. Given the nature of the making activity, manipulation was mainly visually guided. Because of the way in which gaze direction was coded, the categories indicate not only visual attention (what they are looking at) but can also be used as an index of the overall and behavioural states of the participants. An independent coder blind to the hypotheses scored 20% of the videos, and inter-observer reliability was high: Cohen's kappa = .82, 95% CI [.81, .83].

Table 6-3. Behavioural categories used in the study. F: facilitator; L: learner.

Modality	Participant	Label	Category definition
Manipulation	Facilitator	F_manip	F manipulates materials alone
		F_joint_manip	F manipulates materials jointly with L
		F_no_manip	F does not manipulate materials
	Learner	L_manip	L manipulates materials alone
		L_joint_manip	L manipulates materials jointly with F
		L_no_manip	L does not manipulate materials
Vocalization	Facilitator	F_voc	F vocalizes
		F_no_voc	F does not vocalize
	Learner	L_voc	L vocalizes
		L_no_voc	L does not vocalize
Gaze direction	Facilitator	F_look_F	F gazes at her own manipulations
		F_look_L	F gazes at L's manipulations
		F_look_joint_manip	F gazes at their joint manipulation
		F_look_Leyes	F gazes at L's eyes
		F_look_other	F gazes somewhere else or NA
	Learner	L_look_F	L gazes at her own manipulations
		L_look_L	L gazes at F's manipulations

	L_look_joint_manip	L gazes at their joint manipulation
	L_look_Feyes	L gazes at the F's eyes
	L_look_other	L gazes somewhere else or NA

6.2.7 Analysis

I examined the performance of the potential learners before and after the making activity (or distraction task) and the time series representing the participants' manipulation, vocalization, and direction of gaze. With regard to manipulation and vocalization, I simply computed the proportion of time spent in each behavioural category. The time series of gaze direction was used to examine behavioural coordination in detail as described below.

Cross Recurrence Quantification Analysis

To examine the coordinative dynamics of the dyads' joint visual attention in more detail I used a family of nonlinear techniques from recurrence analysis based on cross recurrence plots (**CRPs**). These analyses were conducted using the `crqa` package in `R` (Coco & Dale, 2014). In general terms, the steps to produce a CRP from a pair of categorical time series are the following. First the cross recurrence matrix is computed, allowing only a set of chosen pairs of behavioural states to count towards cross recurrence. In this matrix, columns represent the time indices of the facilitator's time series, and rows those of the child's. The states of the two systems are compared for each combination of time indices. For one-dimensional categorical data such as behavioural states, this comparison is straightforward: if the two participants exhibit 'the same' behavioural category, a value of one is included in the corresponding entry in the matrix, thus indicating cross recurrence; otherwise a value of zero indicates the states are different, i.e. no cross recurrence. **Table 6-4** shows the pairs of behavioural categories considered as being 'the same behavioural state' and which were thus allowed to count towards cross recurrence.

Table 6-4. The pairs of behavioural categories considered to be 'the same behavioural state' for the purpose of cross recurrence quantification analysis.

Behavioural categories		Justification
Facilitator	Learner	
F_look_F	L_look_F	Both are looking at the facilitator's actions
F_look_L	L_look_L	Both are looking at the learner's actions
F_look_joint_manip	L_look_joint_manip	Both are looking at their joint actions
F_look_Leyes	L_look_Feyes	Both are looking at each other's eyes

The cross recurrence matrix can then be used for cross recurrence quantification analysis (CRQA) and can also be plotted as a cross recurrence plot (CRP) for visualization. In the CRP the horizontal axis represents the time indexes of the facilitator's time series and the vertical axis those of the child's, and different colours such as black and white are used to contrast the values of one (i.e. cross recurrence) and zero (no cross recurrence) at each entry in the matrix. As an illustrative example, **Figure 6-4** shows the time series obtained for dyad 7 and **Figure** 6-5 shows the corresponding CRP produced for their gaze direction data in the entire activity. In the CRP, the main diagonal or Line of Synchrony (slope = 1 and passing through the origin) compares the states exhibited by the systems under analysis – here, the two participants – at exactly the same time. Diagonals with slope = 1 and parallel to (above or below) the main compares the states of the participants at different time lags.

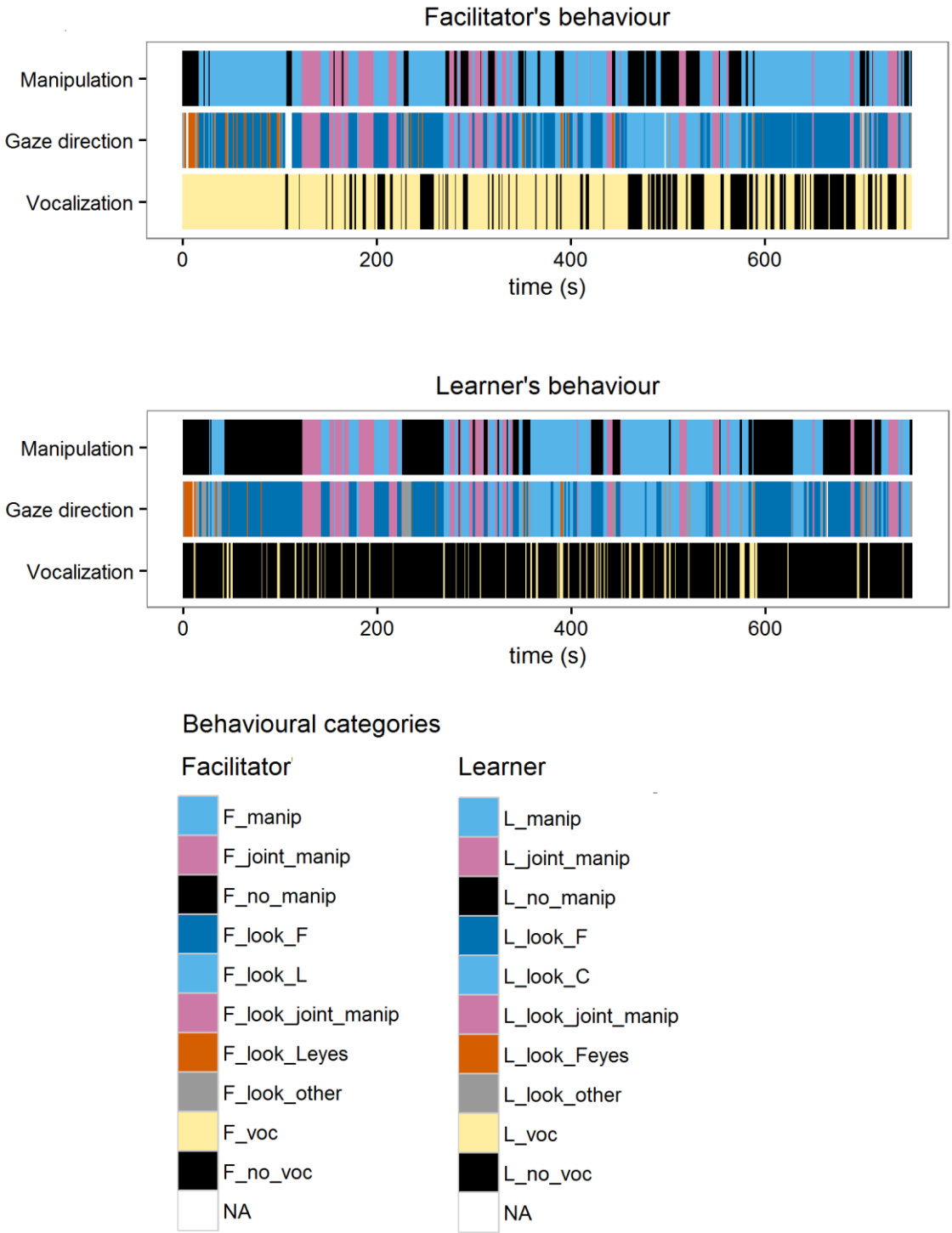


Figure 6-4. Time series coded for dyad 7. The behavioural categories are described in **Table 6-3**.

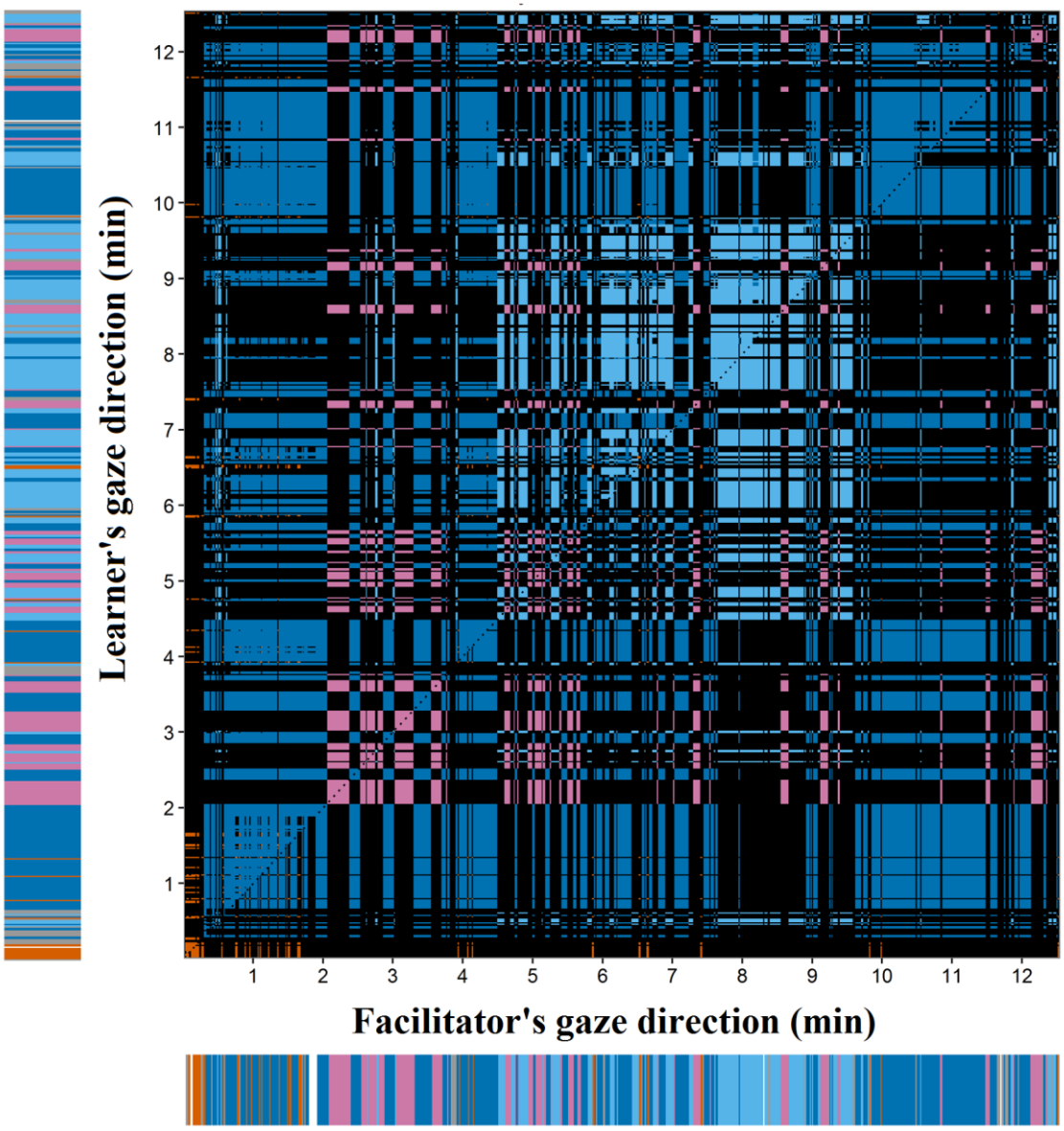


Figure 6-5. Cross recurrence plot (CRP) computed for the gaze direction data from dyad 7. The time series representing the gaze direction of the facilitator (horizontal) and the learner (vertical) are the same as shown in **Figure 6-4**. The CRP has been colour-coded for didactic purpose. Entries in the cross recurrence matrix with a value of zero (no cross recurrence) are in black. Entries with a value of 1 (cross recurrence) are represented with the colour which indicates the corresponding behavioural states.

Cross recurrence points that form continuous lines are of especial interest for quantification analysis (Marwan et al., 2007). Vertical (or horizontal) lines indicate that one (or the other) system is stationary, i.e. that the behavioural state of one (or the other) participant does not change for some time. Diagonal lines with slope = 1 indicate that the trajectories of both systems in their corresponding state spaces are running in parallel at the corresponding time lag. In the context of this study diagonal lines indicate that both participants were visually tracking the same situation or sequence of situations at the corresponding time lag.

Different textures that obtain in the CRP – i.e. how cross recurrence points are distributed – indicate features of the underlying dynamics that can be quantified by computing CRQA variables. Here I focus on four variables: cross recurrence rate (**RR**), determinism (**DET**), mean line length (**L**), and maximum line length (**MAXL**). The variable RR measures the proportion or density of cross recurrence points in the CRP. It is a general and indiscriminate measure of the degree of shared activity or coordination between the signals, corresponding to the more widely known cross correlation sum (Kantz, 1994). The variable DET is defined as the percentage of cross recurrence points that have a common trajectory in the phase space. In the CRP this corresponds to points that form diagonal lines larger than some defined threshold value and is therefore an index of synchronization. L is defined as the average length of diagonals above the threshold and is an index of the proportion of time in which both systems stay synchronized. MAXL is defined as the longest diagonal line present in the CRP; it quantifies the dynamic's robustness to perturbation and is therefore an index of the stability of coupling. The variables DET, L, and MAXL depend on the threshold value which was here set to the minimum of 2. This corresponds to 2 seconds in the original data, thus these variables consider the cases in which participants are in the same behavioural state for 2 seconds or longer. CRQA variables are descriptive in nature and therefore comparisons among conditions, participants, or appropriate baselines are required to draw inferences and examine specific predictions (Marwan et al., 2007; Shockley et al., 2002).

1) Comparing subtasks (Prediction C1)

To investigate if coupled attention is constrained by the emerging task I examined whether different subtasks produced systematically different quantifications of coordination (Prediction 1.2). To this end, coordination variables were computed for each subtask separately. Participants rarely touched any materials during the first subtask (PICK_BODY) so it was ignored. In addition, the two subtasks involving the novel tool (TOOL_DEMO and TOOL_USE) were considered together (as MAKE_WHEELS). Thus, I compared the subtasks MAKE-WHEELS, ASSEMBLE, and DECORATE. It is possible to extract coordination variables from the entire CRPs corresponding to each subtask, as described above. However, doing this would include many lags which are far away from the main diagonal and thus do not seem relevant in this case. For example, for the purpose of examining the on-going coordination between participants, it does not seem relevant to include lags which tells us that the facilitator looked where the child was looking five or ten minutes before. Therefore, it makes sense to focus on a narrow window of a few seconds around the main diagonal. This can be done by conducting a windowed analysis. A windowed analysis consists in building CRPs and computing CRQA measures, not once for the entire time series but multiple times for a sequence of overlapping windows of a determined size, thus effectively tracking how CRQA values evolve over the entire time course. In the CRP shown in **Figure 6-5**, this would correspond to computing CRQA variables, not for the entire CRP but for a set of overlapping windows around the main diagonal. I used the `wincrqa` function and set the parameter `windowsize` to 11 and parameter `windowstep` to 1. This corresponds to using a window including the main diagonal ± 5 seconds lag in each side, computing the coordination variable, moving the window one second, computing the variables again, and so on, for the entire time series. This produces high-resolution data. Thus, I simply computed the average values for each subtask. The values of RR computed for the subtask MAKE_WHEELS (RR_{MW}) were additionally used as a predictor in the analyses of the association between coordination and implicit learning (see below). Finally, I also computed RR for the entire activity ($RR_{activity}$) and this was used as a predictor in the analysis of the association between coordination and explicit learning (see below).

2) Quantifying synchronization (Prediction C2)

To examine the degree of synchronization in coupled attention (*Prediction C2*), diagonal-wise cross recurrence profiles (**DCRP**) were produced for each dyad (Marwan et al., 2007). One example is shown with the filled circles in **Figure 6-6**. The DCRP is produced by plotting the proportion of cross recurrence points that obtain at each lag within a range of interest, here defined to include the main diagonal (lag 0) and 10s on either side. The patterns that obtain in the DCRP indicate aspects of synchronization. If one participant tends to lead the other's attention, this will produce a peak on either side depending on who is leading. In contrast, if both participants take turns leading and following, or if they tend to track the same situation at the same time, then this will produce a peak near lag zero. If there is no clear leader/follower pattern, this will produce a relatively flat curve.

3) Comparing temporal patterns of coupled attention (Prediction C3)

To examine if coupled attention follows a similar temporal organization across dyads (*Prediction C3*), I conducted windowed analyses separately for each relevant pair of behavioural states:

(1) looking at facilitator's actions (categories F_look_F and L_look_F)
(2) looking at children's actions (categories F_look_L and L_look_L)
(3) looking at joint manipulation (categories F_look_joint_manip and L_look_joint_manip)
(4) mutual gaze (categories F_look_Leyes and L_look_Feyes)

As above, the window included the main diagonal ± 5 seconds lag in each side and it moved only one second in each iteration. I report only the variable RR as a general quantification of coupled attention. For this analysis I used the complete, high-resolution data (rather than computing the average, as above) to examine how the dyad's attention changed in real time throughout the entire task. Since dyads varied in the amount of time spent in each subtask, the time series were normalized to be of the same length prior to this analysis. This was accomplished by making the length of each subtask equal to the minimum length observed in the sample, producing time series with a normalized total length of 464 seconds or 7 min: 44 sec. The subtask PICK_BODY corresponded to only 9 secs after normalizing and therefore it is not captured properly in this analysis. These data were not modelled formally but used for qualitative interpretation.

Statistical analysis

All analyses were carried out in `R` version 3.1.3 (R Core Team, 2016). Regression models were fit in STAN (Carpenter et al., in press) using the `stan` function from the `rstan` package (Stan Development Team, 2016) or the convenient `map2stan` function from the `rethinking` package (McElreath, 2016). STAN is a tool for implementing Bayesian multilevel models using Markov Chain Monte Carlo (MCMC) simulation, which approximates the posterior distribution for unknown parameters (Gelman et al., 2014; an accessible introduction is McElreath, 2016). Fitted Bayesian models provide samples from the posterior distribution of the parameters for which summary statistics are computed and reported (such as mean and standard deviation). In all models, I used weakly informative priors to avoid unreasonable parameter values while still allowing the model to estimate a wide range of sensible values (Gelman et al., 2014), and I report posterior distributions estimated from 3,000 samples after warmup.

1) Analysis of coordination across subtasks (Prediction C1)

To investigate if coupled attention is constrained by the emerging task, I modelled the coordination variables computed for each subtask using multilevel linear models. Four models were produced, one for each variable. In each model the outcome was the respective variable and the predictors were an indicator for subtask (MAKE-WHEELS, ASSEMBLE, DECORATE) and included varying intercepts (also called random effects) for dyad.

2) Analysis of synchronization (Prediction C2)

To examine synchronization quantitatively, I summarized the data from all dyads (N=17) with a single test statistic and used a simulation-based approach: I simulated data under two opposing assumptions, summarized each simulation with the same test statistic, and replicated this procedure 10,000 times. I then compared the two resulting distributions of the test statistic computed for the simulations with the test statistic computed from the observed data. The test statistic used to formalize this check was defined as the difference between the maximum RR and the average RR in the profile:

$$\mathrm{diffRR}_i = \max(\mathrm{RR}_i) - \mathrm{mean}(\mathrm{RR}_i)$$

I used this variable (**diffRR**) as a proxy for how peaked each profile was and thus as a measure of the amount of synchronization between participants. **Figure 6-6** illustrates this procedure with an example from a single dyad for didactic purposes. However, for the actual test I used the average obtained across all dyads. The two opposing assumptions tested were that (1) the attentional flow of the participants is independent from each other; and (2) the attentional flow of the participants is perfectly synchronized with each other. By simulating data under the two opposing assumptions, it is possible to examine where the observed data lie between the theoretical extremes. For this analysis I allowed only the following pairs of categories to contribute towards cross recurrence: F_look_F and L_look_F; F_look_L and L_look_L. By excluding the categories F_look_joint_manip and L_look_joint_manip I am choosing for a conservative approach, since including them would inevitably increase the amount of synchronization detected because of how they were defined (described in **Table 6-3**).

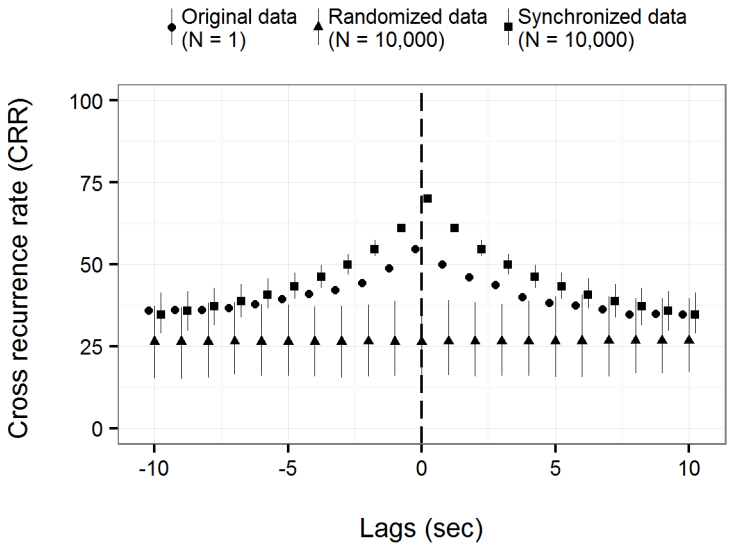


Figure 6-6. Diagonal-wise cross recurrence profile computed for dyad 7, showing the original data (filled circles) as well as two sets of 10,000 simulated data corresponding to the randomized (triangles) and perfectly synchronized (squares) baselines. Lag zero correspond to the main diagonal in **Figure 6-5**, and the ± 10 sec lags correspond to diagonals above/below it. To avoid overplotting and improve readability, the original data were shifted slightly to the left and the synchronized data to the right.

3) Analysis of learning (Predictions L1 and L2)

To examine if children's performance with the tool improved as a result of engaging in the activity (*Prediction L1*) I modelled the data from the implicit knowledge tests using multilevel logistic regression with logit link. The outcome was the probability to succeed

with the tool (i.e. to produce a hole) and the predictors were an indicator for treatment (control = 0, making = 1), an indicator for test (Test-1 = 0, Test-2 = 1) and the interaction between them.

To examine whether children's ability to identify the steps in the task (explicit knowledge) improved from Test-1 to Test-2, I modelled the data from the explicit knowledge tests using multilevel ordered categorical regression with latent variable and cutpoints (following Gelman & Hill, 2007; and implemented according to McElreath, 2016). The outcome was the cumulative probability to obtain each possible score in the explicit knowledge tests (i.e. the number of correct subtasks spontaneously identified by the children, either 0, 1, 2, 3, or 4). The predictors were an indicator for test (Test-1 = 0, Test-2 = 1) and an error term for children to account for individual variability. Dyads 8, 10, and 11 were excluded due to experimenter error during the tests, therefore this analysis included N=14 sessions. To specifically test whether children scored higher in the post-activity test compared to the pre-activity test, I computed the posterior difference between the corresponding coefficients for Test-2 and Test-1.

4) The association between coordination and learning (Prediction L3)

To examine if higher values of attentional coordination were associated with an increase in the ability to use the novel tool, I focused on the subset of children who engaged in the making activity *and* used the novel tool successfully in Test-2 (N = 14). The variable used as an index of coordination was RR_{MW}, the average cross recurrence rate computed during MAKE-WHEELS, as explained above. This variable seems appropriate as it captures the attentional coordination between children and the facilitator during the only subtask which involved the hand drill. Three generalized linear models with gamma error structure and logarithmic link were fitted to the data and compared in terms of their prediction accuracy using WAIC scores (Widely Applicable Information Criterion; see McElreath, 2016). Lower values of WAIC indicate better accuracy. In all models the outcome was the time children required to produce a hole in Test-2 (timeT2). The first model had only one predictor, RR_{MW}, which was centred (i.e., subtracted from the mean) to facilitate interpretation. The second model included as a second predictor the children's centred score in the explicit knowledge Test-1 (EKTest-1). This corresponds to statistically controlling for the children's existing motor abilities prior to

the activity. A better control would have been the time children required to produce a hole in Test-1, but no child succeeded in the initial test. The third model included the second-order interaction between RR_{MW} and EKTest-1.

To examine whether higher values of coordination during the activity were associated with more steps being explicitly identified in Test-2 than Test-1, I modelled the data from the explicit knowledge tests using multilevel ordered categorical regression with latent variable and cutpoints (Gelman & Hill, 2007). Three models were produced and compared using WAIC. In all models the outcome was the cumulative probability to obtain each possible score in the explicit knowledge Test-2. The first model had only one predictor, the cross recurrence rate computed for the whole activity ($RR_{activity}$), which was centred prior to including in the model. The second model included the centred score obtained in Test-1 as a covariate (EKTest-1) to adjust for each participant's baseline score. The third model included the second-order interaction between $RR_{activity}$ and EKTest-1. As with Prediction 2.3 above, this analysis included N=14 sessions.

Evaluating evidence of effects

All STAN models converged, and mixing was good. Unless otherwise indicated, I report the mean and 95% credible interval of the estimated parameters from the fitted models. In this section, strong evidence for an effect corresponds to a central credible interval of at least 95% which excludes 0, and weak evidence corresponds to cases where the 95% central credible interval includes 0, but the 90% central credible interval does not. I sometimes report the probability that the model estimates a positive parameter value to provide evidence of an effect.

6.3 Results

6.3.1 General characterization of the activity

With regard to the manipulative actions, on average participants manipulated materials alone half of the time (facilitator: 47%, range [31%, 61%], children: 53%, range [37%, 75%]) and there was no evidence of a systematic difference between facilitator and children (Welch's t test: 95 % CI of the difference in seconds [-15, 103]). In addition, on average, dyads jointly manipulated the same materials during 17% of the time (range [8%, 26%]). The facilitator vocalized on average during 52% of the time (range [38%, 69%])

and children much less, during only 9% of the time (range [2%, 16%]) and there was strong evidence that this difference was systematic (Welch's t test: 95 % CI of the difference in seconds [263, 349]) and corresponded to a large effect size (r = .95). Qualitative assessment of the recordings suggest that the facilitator used speech in a variety of contexts, including to engage children's imagination ("*What sort of vehicle do you want to make?*"); to indicate how she anticipated the situation would unfold and thus possibly influencing the children's expectation ("*I'm going to show you how we're going to use the drill, and then you're going to have a go, ok?*"); to narrate her actions during the hand drill demonstration and thus possibly facilitate learning by directing children's attention to specific aspects of the unfolding task ("*You're going to hold the top of this [handle] in your left hand…*"); to provide feedback and instruction ("*That's it!*", "*Keep turning!*"), to direct children's attention to specific materials ("*There's tissue paper, stars, plasticine…*"), and to finish the activity and praise the children's work ("*Right, are we done? That is a brilliant ice cream van!*"). Children, on the other hand, mainly used vocalization to provide backchannel responses ("*Aham*", "*Yeah*"), and to express emotions such as surprise ("*Wow*") and frustration ("*Oh*").

Visual attention will be investigated in more depth below, but as a first step I operationalized synchronized coupled attention simply as the proportion of time in which both participants were visually tracking the same target situation during the same second interval – which target could be either the child's hands, the facilitator's hands, or their joint manipulation. With this preliminary definition, dyads engaged in synchronized attention on average for 74% of the time (range [68%, 81%]). These general patterns are overall consistent with the task being a joint making activity in which participants played different and complementary roles, sometimes doing things together and sometimes alone, and in which the facilitator used her actions and vocalizations to influence the behavioural flow of the children with the intention to follow the pre-defined script.

6.3.2 Analysis of coordination across subtasks (Prediction C1)

I predicted different subtasks would produce systematically different signatures of attentional coordination. **Figure 6-7** shows the model estimates for each of the four coordination variables (CRR, DET, L, MAXL). Overall, estimated CRQA values were similar across subtasks, indicated by their broadly overlapping ranges. However, there

was evidence of systematic differences across subtasks, indicated by the peaked pattern in all plots. **Table 6-5** shows the probability that the models estimated the highest and lowest values for each subtask. With regard to the subtask ASSEMBLE, there was strong evidence that the dyads showed the highest values of overall coordination compared to the other subtasks (RR: p = .96), that they synchronized more (DET: p = .99), and that they stayed synchronized for longer (L: p = .97). In addition, with regard to the subtask DECORATE, there was strong evidence that they were less synchronized (DET: p = .94) and that the behavioural coupling was less stable (MAXL: p = 1).

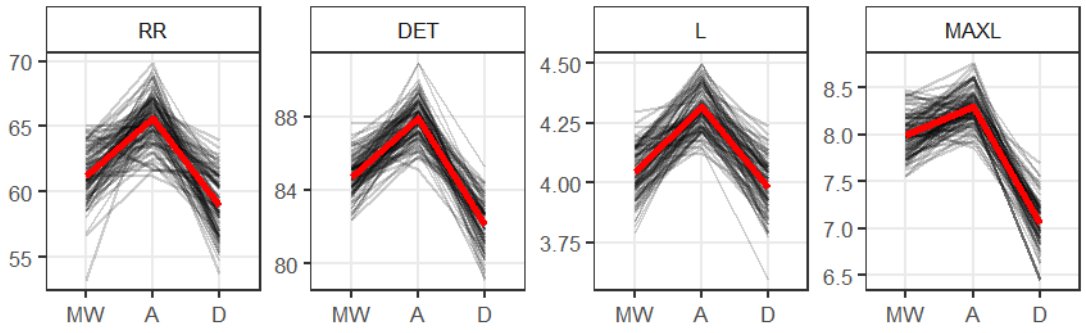


Figure 6-7. Estimated values of the coordination variables RR, DET, L, and MAXL, in subtasks MAKE_WHEELS (MW), ASSEMBLE (A), and DECORATE (D). Reported values are the mean estimates (thick, red) and 100 simulations from the fitted models (thin, black). Even though the values overlap across subtasks, it is possible to see that the models systematically estimated the highest values for the subtask ASSEMBLE.

These results are consistent with the prediction that the different constraints imposed by each subtask should produce systematic differences across conditions. On the one hand, during ASSEMBLE, the subgoal was to have the basic vehicle ready by bringing the set of materials available at that stage – the body, four wheels, and four axles – into a specific spatial arrangement. The manipulative behaviours required to transform their current situation were constrained by the properties of the materials and the imagined design. They were also constrained by the children's existing manual dexterity, which is a function of their prior history in similar making activities. The facilitator must be aware of all these elements and intervene appropriately, sometimes even taking full control of the flow of manipulations (i.e., going from manipulating jointly to manipulating alone) without fully disengaging the children from the activity altogether. The results indicate that these challenges were particularly demanding in terms of behavioural coordination. Qualitative assessment of the videos was consistent with this: children struggled to

complete the subtask without assistance and the facilitator attended closely to the unfolding of events and intervened multiple times.

On the other hand, during DECORATE, instead of the rigid normative rules imposed by the script, children were relatively free to choose among a set of familiar materials **Figure 6-2** and were invited to decorate the vehicle 'however they wanted'. These constraints produced less behavioural coordination between the facilitator and the children. Qualitative assessment of the videos corroborated this: the facilitator frequently moved away and looked away from the children's hands more often, as children explored the different materials. However, the facilitator still supervised, suggested alternatives, assisted, encouraged, and praised them until she eventually started bringing the activity to an end.

Table 6-5. Probability that the models estimated the highest or the lowest value of coordination for each of the subtasks. For example, the model of RR estimated that the subtask MAKE-WHEELS will have the highest value with probability .06 and the lowest value with probability .24. Probabilities lower than .1 or higher than .9 are in bold. The cases reported in the main text are highlighted in grey.

Variable	Subtask					
	MAKE-WHEELS		ASSEMBLE		DECORATE	
	p(highest)	p(lowest)	p(highest)	p(lowest)	p(highest)	p(lowest)
RR	**.04**	.21	**.96**	**0**	**0**	.79
DET	**.01**	**.06**	**.99**	**0**	**0**	**.94**
L	**.02**	.28	**.97**	**0**	0	.72
MAXL	.15	0	.85	**0**	**0**	**1**

6.3.3 Analysis of synchronization (Prediction C2)

I predicted participants' visual attention should be synchronized in addition to being coordinated. This was examined with the diagonal-wise cross recurrence profiles. In all dyads the maximum RR was observed at lag zero, supporting the prediction, i.e. participants tended to be looking at the same situation *at the same time*. Furthermore, there was no clear leader/follower pattern, indicating that the facilitator was not only guiding the children but also being guided by the children. This is constrained by the resolution available (1 observation per second) and it should be noted that a finer resolution might produce a richer result.

The test statistic *diffRR*, defined above, was interpreted as an index of synchronization. I compared the mean value of *diffRR* computed for the observed data ($diffRR_{observed}$) with the distribution of values computed for randomized simulated baseline ($diffRR_{random}$) and the perfectly synchronized simulated baseline ($diffRR_{synced}$). **Figure 6-8** shows the distribution of *diffRR* values computed for the simulated baselines and the value computed for the observed data. There was strong evidence that the observed data were indeed synchronized (probability that $diffRR_{random}$ is equal to or higher than the value of $diffRR_{observed}$, $p = 0$). This is consistent with the suggestion that, in this case, behavioural coupling is direct (participants influence each other in real time) and the influence is bidirectional (the facilitator influences the children and the children also influence the facilitator). In addition, there was strong evidence that this synchronization was not perfect (probability that $diffRR_{synced}$ is equal to or higher than the value of $diffRR_{observed}$, $p = 1$). This is consistent with participants influencing each other's behaviour in the joint making activity while maintaining their individual autonomy.

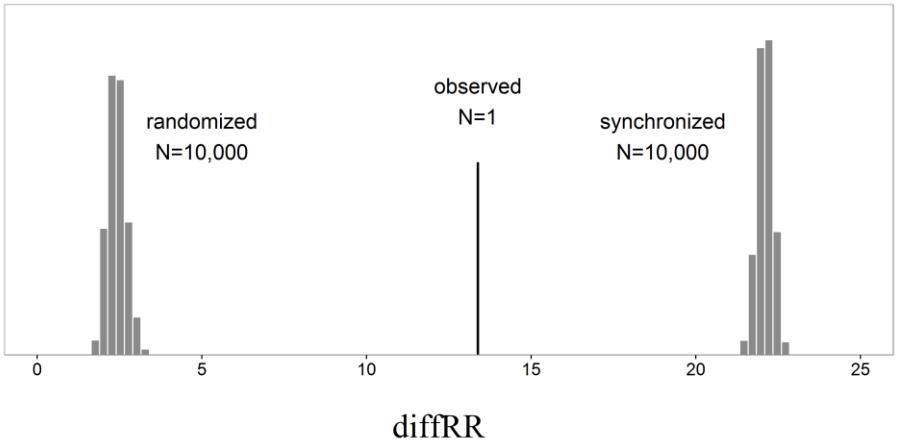


Figure 6-8. Distribution of the test statistic diffRR computed for the two simulated baselines, randomized and perfectly synchronized. The solid vertical line shows the value of diffRR computed for the observed data.

6.3.4 Temporal patterns of coupled attention (Prediction C3)

I expected different dyads would produce similar temporal patterns of coupled attention because the facilitator was explicitly committed to follow the same script every time. **Figure 6-9** shows the median results from the windowed analysis, and **Figure 6-10** supplements it by representing variability across dyads. In these plots the horizontal axes represent time as the sequence of overlapping windows produced in the windowed analysis, and the vertical axes represent the RR values computed in each window for each

target of interest (both looking at the facilitator's actions, both looking at the children's actions, both looking at their joint manipulation, mutual gaze). Relatively higher values of RR indicate that the corresponding target dominates the dyad's coupled attention. Therefore, these results capture the dynamical constitution of coupled attention and provide evidence to support a narrative about how coupled attention is established, maintained, and eventually dissolved as the task unfolds.

The proportion of time participants spent looked at each other's eyes was overall very small, but the facilitator gazed at the children relatively much more (mean: 6%, range [4%, 8%]) than children gazed at the facilitator (mean: 3%, range [1%, 4%]). This is reflected in extremely low RR values for the cross recurrence of the categories [F_look_Leyes and L_look_Feyes] throughout the task, with the notable exception of the initial moments, where most of the mutual gaze occurred. During the subtask MAKE-WHEELS, the dyads' coupled attention was first directed to the facilitator's hands and then to the children's or to their joint manipulation. This is consistent with the facilitator first demonstrating the tool (TOOL_DEMO) and then the children using it four times to make the wheels (TOOL_USE). Note there are four waves of 'both looking at their joint manipulation', corresponding to the children using the hand drill four times to produce four wheels. These curves have decreasing peaks, suggesting children overall were faster and/or received less assistance from the facilitator as they accumulated practical experience with the tool. During ASSEMBLE, the dyads' coupled attention was again first directed to the facilitator's hands and then to the children's or to their joint action. This is consistent with the facilitator first prompting the children and then assisting them put the parts together. Finally, during DECORATE, the dyads' coupled attention was first directed to the facilitator's hands and then shifted mostly to the children's. This is consistent with the facilitator first prompting the children and then the children working creatively with the common materials to decorate their vehicle, thus receiving less guidance. **Figure 6-10** tells the same general story but indicates the variability across dyads.

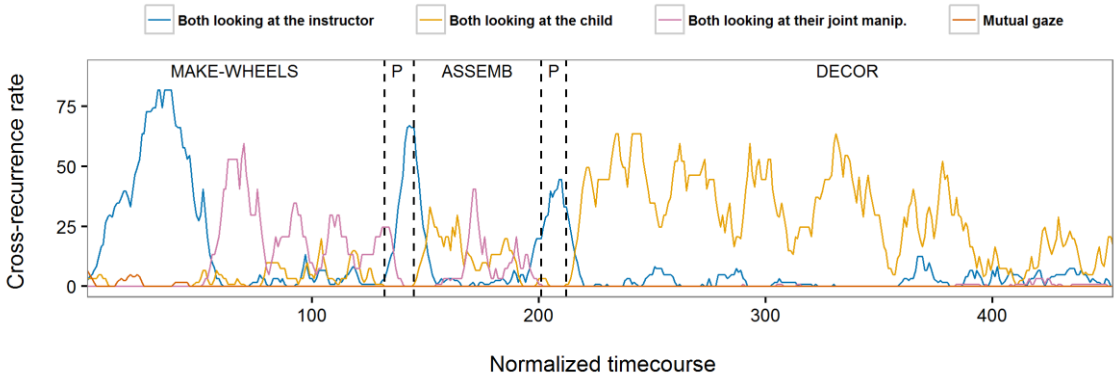


Figure 6-9. Temporal evolution of the targets of coupled attention as the task unfolded. Windowed analyses were conducted separately for each target of attention after normalizing the length of the time series as explained in the main text. The coloured lines show median values of cross recurrence rate (across all dyads) at each window, for each of the possible targets of attention. The dashed vertical lines indicate the transitions between subtasks. The time intervals P have been added to indicate when prompts were being given by the facilitator.

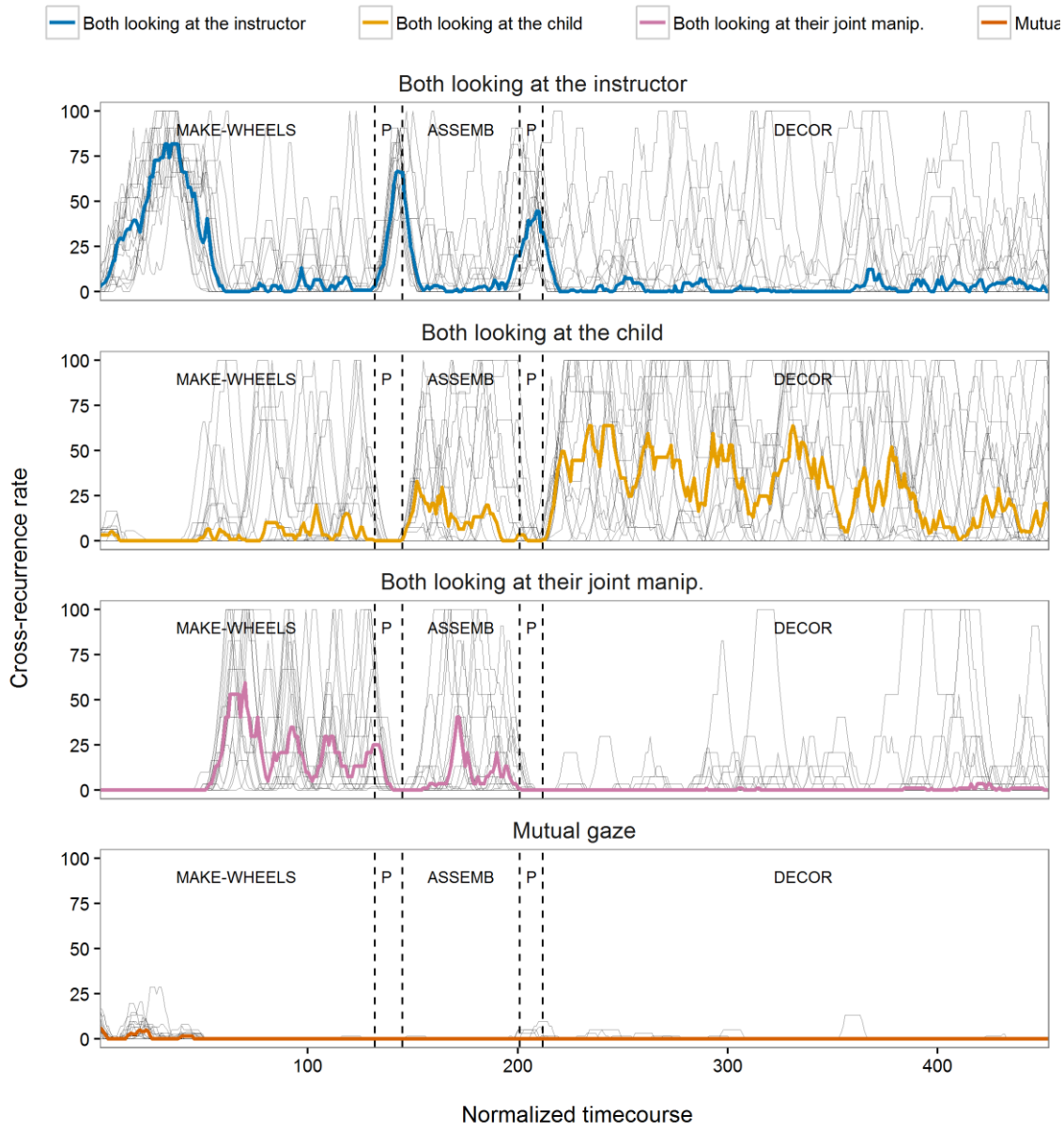


Figure 6-10. Temporal evolution of the targets of coupled attention as the task unfolded. Windowed analyses were conducted separately for each target of attention after normalizing the time series. The thick coloured lines are the same as in **Figure 6-9** and show median values of cross recurrence rate (across all dyads) at each window, for each of the possible targets of attention. The black lines show the data for each of the N=17 dyads to indicate the variability in the complete dataset. The dashed vertical lines indicate the transitions between subtasks. The time intervals P have been added to indicate when prompts were being given by the facilitator.

6.3.5 Analysis of learning (Predictions L1 and L2)

All children responded that they had seen a pair of scissors before, knew scissors are for cutting, and had used them before, thus suggesting they understood this set of questions about familiarity. In contrast, only two children in the control group and two in the experimental group reported having seen a hand drill before, and all children reported that they had never used a hand drill before. I therefore considered the hand drill to be a novel tool for all participants.

Implicit learning

I predicted that children who engaged in the making activity would improve their ability to use the novel tool more, compared to children who used the tool the same number of times but unassisted. **Figure 6-11** reports the fitted logistic model. The estimated mean probability to succeed before the activity (Test-1) was similar for children in both conditions and there was no evidence of a difference between conditions (estimated difference in the probability to succeed: .10, 97% CI [-.01, .27). This suggests that children started with a roughly similar ability in both conditions. In contrast, children were much more likely to succeed after (Test-2) the joint making activity than after the distraction activity and there was strong evidence of a positive difference between conditions (estimated difference in the probability to succeed: .62, 97% CI [.32, .83]). The results indicate that engaging in the making activity with the adult maker indeed facilitated learning, as intended.

Explicit learning

Figure 6-12 reports the model of explicit knowledge predicted by test. The uncertainty around the parameters is very broad. However, as predicted, there was strong evidence that children explicitly identified more steps after the activity compared to before (difference between $b_{Test\text{-}2}$ and $b_{Test\text{-}1}$: 1.58, 95% C.I. [0.01, 3.36]). This indicates that participating in the making activity hands-on changed, indeed improved, children's ability to imagine and explicitly identify the steps involved in making a toy vehicle.

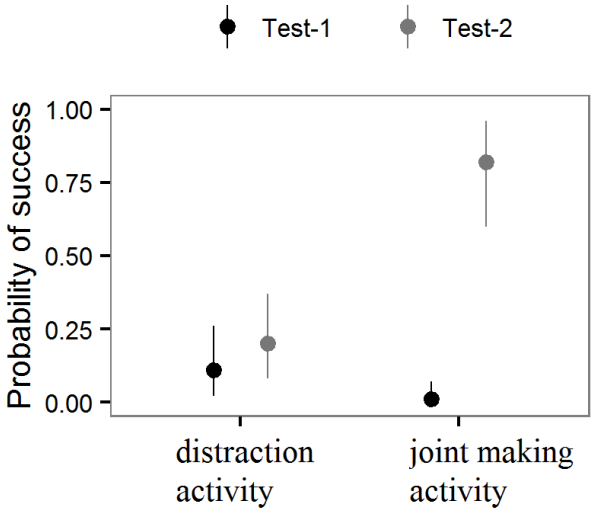


Figure 6-11. Evidence of implicit learning. Estimated probability to succeed using the novel tool before (Test-1) and after (Test-2) participating in the distraction activity or the joint making activity.

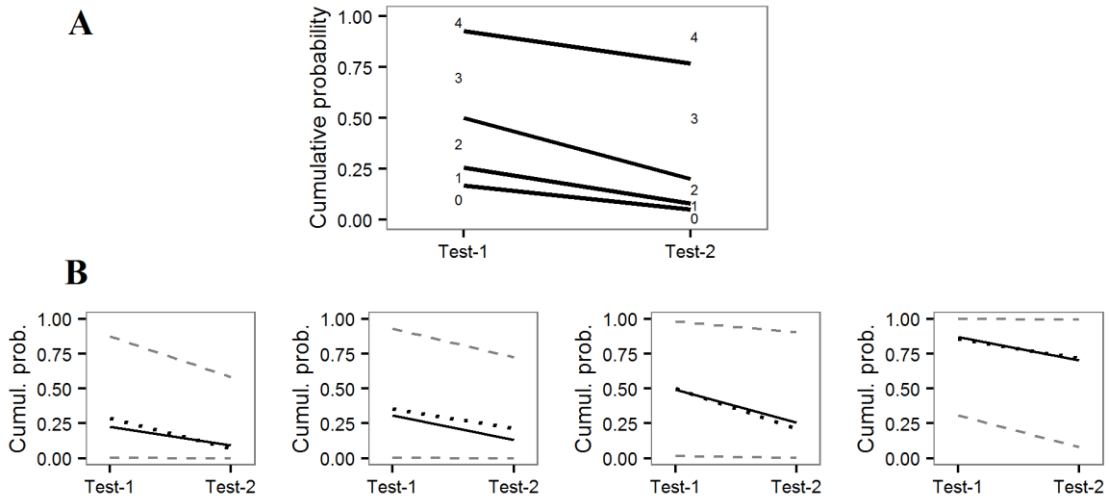


Figure 6-12. Evidence of explicit learning. A) Mean estimated cutpoints indicating the cumulative probability to identify 0, 1, 2, 3, or 4 steps in the explicit knowledge tests, before (Test-1) and after (Test-2) participating in the joint making activity. For example, the third cutpoint from the bottom indicates that the mean estimated probability to identify 3 or 4 steps was about .5 in Test-1 and about .75 in Test-2. B) The mean (black solid lines, same as in A) and 95% credible intervals (grey dotted lines) of the estimated cutpoints are here shown separately to indicate model uncertainty.

6.3.6 The association between coordination and learning (Prediction L3)

Implicit learning

I predicted that higher values of coordination during the subtasks involving the novel tool (MAKE-WHEELS) would be associated with more time-efficient tool use in Test-2. The results from model comparison are shown in **Table 6-6**. The interaction model obtained the lowest WAIC score and an Akaike weight of .74, suggesting this the best model given the data and model assumptions. Thus, the decision to include the children's score in the explicit knowledge Test-1 (EKTest-1) seemed to have been appropriate as an index of the children's existing motor abilities prior to the activity.

Table 6-6. Model comparison using information criteria of the three models examining the association between implicit learning and coordination. WAIC: Widely Applicable Information Criterion; dWAIC: difference between each WAIC and the lowest WAIC; weight: Akaike weight (rescaled WAIC); SE: standard error of WAIC estimates; dSE: standard error of the difference in WAIC between each model and the top-ranked model.

Predictors in each model	WAIC	dWAIC	weight	SE	dSE
RR_{MW} + EKTest1 + RR_{MW} X EKTest-1	108.9	0.0	.74	8.9	NA
RR_{MW}	111.7	2.8	.19	11.33	11.96
RR_{MW} + EKTest-1	113.5	4.5	.08	10.86	10.64

Table 6-7 reports the estimated coefficients of the interaction model. For interpretation, **Figure 6-13** shows model predictions as a tryptic, with the scores for the centred predictor EKTest-1 fixed at lowest (-2), mean (0), and high (2) values. For children who indicated having more prior experience in similar activities (higher EKTest-1 scores), coordination during MAKE-WHEELS was not associated with performance with the tool in Test-2. This is indicated by the flat predictive curve of the right plot in **Figure 6-13**. On the other hand, for children who indicated having less prior experience in similar activities (lower EKTest-1 scores), there was strong evidence that coordination during MAKE-WHEELS was associated with performance with the tool in Test-2 but in the opposite direction than predicted. Higher values of RR_{MW} were associated with higher (less efficient) values of time to succeed in Test-2. **Table 6-7** indicates this in the positive effect of RR_{MW} and **Figure 6-13** illustrates this in the ascending curve of the leftmost and middle plots. In a post-hoc analysis, including age as a covariate did not improve model prediction accuracy.

Table 6-7. Summary of the estimated coefficients in the final model examining the association between implicit learning and coordination. N_eff: estimate of the number of independent samples from the posterior distribution, Rhat: a measure of convergence/mixing of independent MCMC simulations. N_eff and Rhat are used as general diagnostics for MCMC models.

Coefficient	Mean	SE	95% C.I.	N_eff	Rhat
Intercept	2.42	0.16	[2.11, 2.75]	728	1
RR_{MW}	0.13	0.04	[0.06, 0.20]	592	1
EKTest-1	0.14	0.14	[-0.13, 0.40]	597	1
RR_{MW} X EKTest-1	-0.06	0.02	[-0.11, -0.02]	607	1
scale	2.92	0.71	[1.57, 4.21]	1409	1

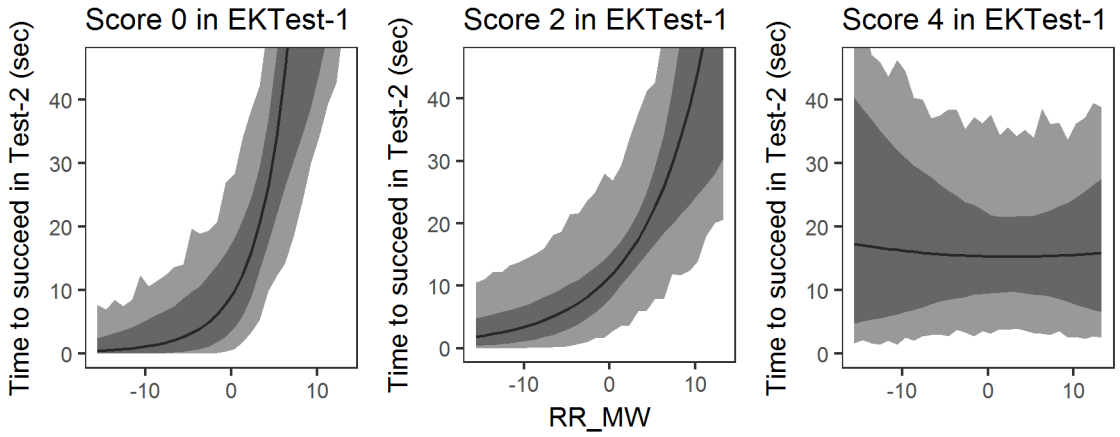


Figure 6-13. Time to succeed with the novel tool after the making activity (Test-2) as a function of coordination during the subtask MAKE-WHEELS (centred RR_{MW}), estimated separately for the lowest (left), average (middle), and highest (right) scores in the explicit knowledge test before the activity.

Explicit learning

I predicted that higher values of coordination during the activity would be associated with children identifying more steps in Test-2 The results from model comparison are shown in **Table 6-8**. The simplest model, including only the measure of coordination in the whole activity (RR) as a predictor, obtained the lowest WAIC score and an Akaike weight of .69, suggesting this is the best model given the data and model assumptions. I therefore report this model.

Table 6-8. Model comparison using information criteria of the three models examining the association between explicit learning and coordination. WAIC: Widely Applicable Information Criterion; dWAIC: difference between each WAIC and the lowest WAIC; weight: Akaike weight (rescaled WAIC); SE: standard error of WAIC estimates; dSE: standard error of the difference in WAIC between each model and the top-ranked model.

Predictors in each model	WAIC	dWAIC	weight	SE	dSE
RR	37.7	0	0.69	12.56	NA
RR + EKTest-1	39.8	2	0.25	12.14	6.08
RR + EKTest-1 + RR * EKTest-1	42.8	5.1	0.05	11.87	7.55

Figure 6-14 reports the chosen model. There was strong evidence that coordination during the activity was associated with explicit learning but in the opposite direction than predicted (b_{RR} = -0.45, [-0.83, -0.12]). Dyads with lower values of coordination during the activity identified more steps after the activity. In contrast, dyads with higher values of coordination during the activity identified fewer steps after the activity. In a post-hoc analysis, I added age as a covariate, but this did not improve model accuracy.

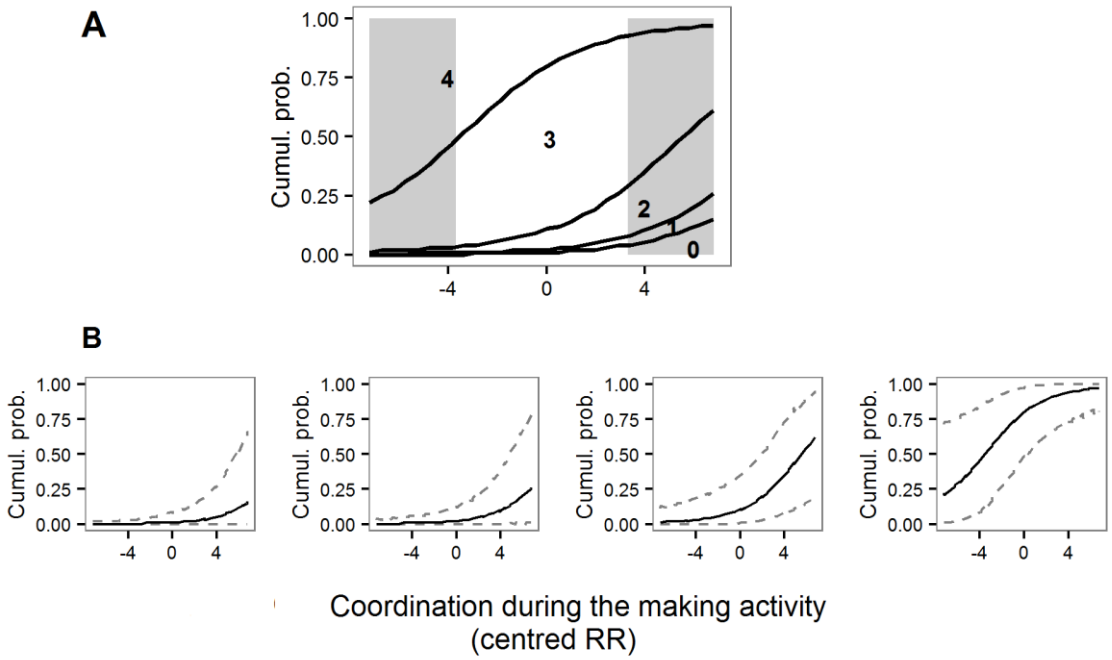


Figure 6-14. Estimated cutpoints indicating the cumulative probability to identify 0, 1, 2, 3, or 4 steps in the explicit knowledge test after the activity (Test-2) as a function of coordination during the whole activity (RR). A) Mean estimates. Shaded areas indicate the 25% lowest (left) and highest (right) values of coordination. B) The mean (black solid lines, same as in A) and 95% credible intervals (grey dotted lines)

of the estimated cutpoints are here shown separately to indicate model uncertainty. For example, children with the lowest 25% values of RR during the whole activity (left shaded region in A) had a mean probability of about .98 to score 3 or 4 in the explicit knowledge Test-2, and a mean probability of only .02 to score 0, 1 or 2. On the other hand, dyads with the highest 25% values of RR (right shaded region in A) had a mean probability of about .55 to score 3 or 4, and a mean probability of .45 to score 0, 1, or 2.

6.4 Discussion

Making and learning together as a relational process

Children were invited to participate in a making activity together with an adult facilitator. This study was used as an opportunity to explore, in more concrete terms, what it means to view social learning as a form-generating process occurring within histories of mutual engagement, rather than in terms of the transmission of information (*sensu* represented knowledge) about the target task. To this end, I focused on the behavioural coordination between the facilitator during the joint making task and the children and its association with learning. **Figure 6-15** represents making and learning together, conceived of as a relational developmental process, schematically.

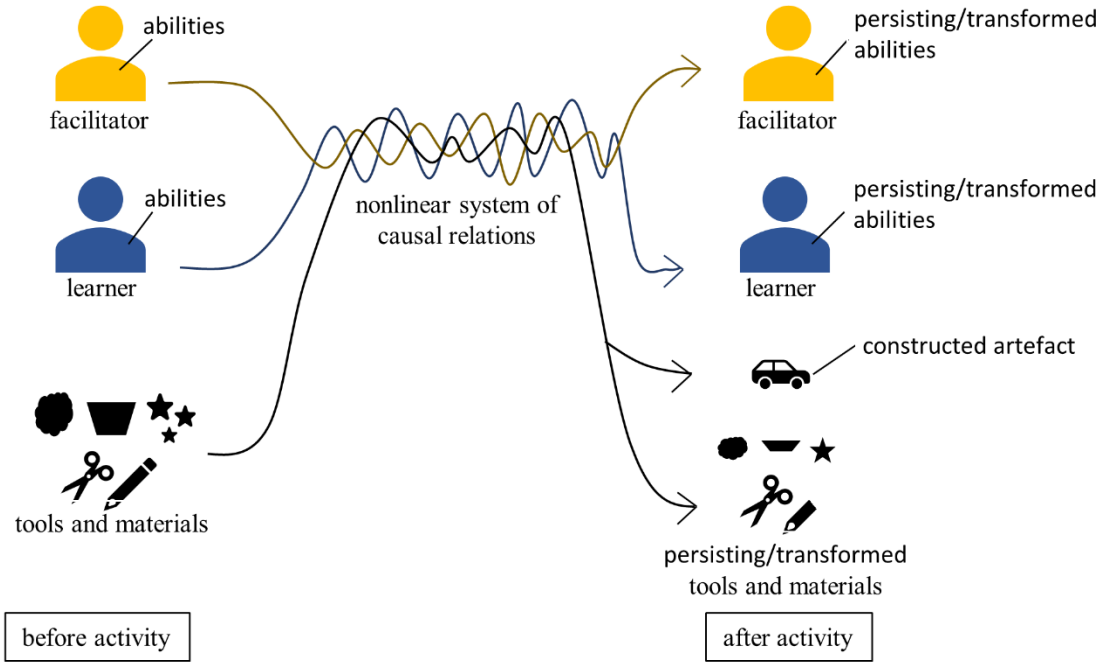


Figure 6-15. Schematic representation of the activity of making and learning together viewed as a relational developmental process.

As the activity unfolded, the tools and materials persisted under transformation and the form of the constructed vehicles appeared as a result of the forces produced through manipulation. Other materials persisted untouched or became rubbish and were discarded.

Similarly, the children's bodies also persisted under transformation, with some aspects remaining the same and others changing. Here I focused on two aspects of what children might learn in the activity. I used the implicit knowledge tests to examine changes in their ability to use the target tool (the hand drill), and I used the explicit knowledge test to examine changes in their ability to imagine and verbally report the steps required to 'solve the task' (i.e., to produce a toy vehicle under the given constraints). In this discussion, I will first focus on behavioural coordination throughout the making activity, then on learning, and finally on the association between coordination and learning.

Making together involves coordinating behaviour in real time

In each session, the child and the facilitator contributed to the making activity by manipulating tools and materials, sometimes alone, sometimes in parallel, and sometimes jointly. However, their roles were not symmetric. The facilitator committed to following the predefined script and her role was to guide the children's behaviour, assist as required, and ensure they reached the goal safely and within the given time frame. In contrast, the children did not know about the script but accepted the invitation to 'make a vehicle' with the facilitator. The children's role was to follow the instructions given by the facilitator and manipulate the tools and materials according to their existing abilities.

Children were not physically forced to do anything. In fact, they were told they could leave at any point without providing a reason (no one did). Thus, to successfully complete the task, the facilitator and the children must coordinate their behaviour according to the contingencies of the unfolding situation and constrained by the task requirements. The fact that all dyads did end up doing similar things, as shown by them producing toy vehicles and enacting similar temporal patterns of coupled attention in each session, is not trivial but noteworthy. This is likely to not have been the case with pre-verbal children, children with atypical neurodevelopmental trajectories and, of course, even less likely if the facilitator were interacting with another species.

It is fair to ask, then, what accounts for the children being able to engage in the activity in similar ways? Part of the answer has to do with the fact that these children have had a similar previous developmental history thus far. For example, their bodies have developed a similar anatomy with similar biomechanical properties (Thelen & Smith, 1994), perceptual systems (J. J. Gibson, 1966), and neuromuscular coordinative structures

(Kelso, 1995). They speak and understand spoken English with the local accent. They all go to schools with similar activities where they pay attention to teachers and follow their instructions. They are familiar with using the tools and materials provided (excepting the hand drill); they have seen vehicles and toy vehicles before and thus have preferences and expectations about what to do.

The similarities in the children's developmental trajectories, in turn, partly depend on the genealogical relations that exist among them as a result of the *evolutionary* history in which their individual ontogenies are nested. Because of their close genealogical relations and the similarities in their ontogenies, their bodies ended up having a similar organization and supporting similar abilities. It is this material similarity and the mutual sensitivity it enables that underlie the way in which the behaviour of the participants might affect each other. One of the main contributions of this study is that it explored a new methodology to capture aspects of the dynamics of this mutual influence.

The dynamical constitution of coupled visual attention.

During the short time they lived together, the behaviour of the participants may become coupled by informational constraints (by means of perceiving each other or when they consistently look at the same targets at the same time) as well as mechanical constraints (e.g., when the facilitator holds the hand drill in place by pushing it down from the top as the child attempts to operate it). This behavioural coupling involved coordinating visual attention with each other and in relation to the tools and materials included in the unfolding situation. Participants established, sustained, and eventually dissolved coupled visual attention constrained by what they were doing, by what the other was doing, and, more generally, by the changing situations in which they found themselves. In other words, coupled attention is not a static, homogeneous state but a historical achievement involving nonlinear causal relations. Cross recurrence quantification analysis captures some dynamical aspects of this process and can therefore be a valuable analytical tool in behavioural studies to complement aggregate measures such as duration and proportion of behavioural categories.

While living – and making – together, the behavioural flows of the participants not only influenced each other but became relatively synchronized, as the facilitator and the children took turns in leading and following each other's attention throughout the activity.

Such synchronisation – i.e., the observation that participants often attended to the same things at the same time, as the activity unfolded – is consistent with the nature of the joint task. The observed synchronisation is consistent with previous research on coupled attention during live conversation (D. C. Richardson, Dale, & Kirkham, 2007). One possible shortcoming of the present study is the relatively coarse resolution in the time series data (one observation per second). Future work could attempt to obtain a finer spatiotemporal resolution to produce a more refined picture of leader/follower patterns than the ones reported here. In the study reported in the next chapter, I used an eye-tracking system to obtain data with higher resolution.

The inclusion of a randomized baseline to test whether participants synchronize their behaviour above chance is common in studies using CRQA (e.g., D. C. Richardson & Dale, 2005; D. C. Richardson et al., 2007). In addition to this, I also simulated a perfectly synchronized baseline. I suggest that simulating both the randomized baseline and the perfectly synchronized baseline is useful because it provides a range, a kind of synchronization scale, within which the observed data could be located. In the present study, this provided evidence that, although the behaviour of participants synchronizes, this synchronization is not perfect. This result is sensible as it indicates that participants maintained their autonomy while at the same time being influenced by the partner's behaviour (De Jaegher & Di Paolo, 2007; De Jaegher & Froese, 2009; Thompson, 2007; Varela et al., 1991). This innovation can be seen as a contribution to methodology in the study of behavioural coordination more generally and specifically in the study of living and learning socially.

Learning socially: not transmission, but development

Before the activity, the facilitator already knows how to make a toy vehicle and how to use the hand drill, whereas the children do not. After the activity, the facilitator still knows these things and now the children do, too, at least to some degree. These empirical observations suggest that the facilitator 'passed on' some form of knowledge to the children during the activity. In a loose sense, talking in these terms is unproblematic. One way to make sense of the observations regarding tool use, following a representational view of cognition and the linear communication model, is by saying the following:

1. Successful tool use requires that the person has relevant information (*sensu* represented knowledge) about how to use the tool.
2. Before the activity, only the facilitator, but not the children, has such information.
3. During the activity, the facilitator transmits this information to the children.
4. After the activity, both the facilitator and the children have this information.

This general picture might be useful for those content with a computational-representational view of cognition. In addition, for studies focusing, not on dyads but on large-scale spatiotemporal patterns spanning across many individuals and across many generations, such as those using mathematical or computational models, it makes good sense to find economic ways to talk about how knowledge may be 'passed on'. However, for those interested in a processual view of the 'passing on' of knowledge in terms of the ontogeny of bodies-in-their-environment, rather than in terms of the execution of computational algorithms or in terms of the evaluation of mathematical models, the 'transmission' metaphor becomes less appealing. By focusing on abstract computational or mathematical events rather than concrete behavioural or physiological ones, the transmission metaphor may drive our attention away from the processes we set out to investigate.

Consider, for example, what happens when the facilitator attempts to demonstrate the correct use of the novel tool to the children (during subtask TOOL_DEMO), and suppose the children are attending to the unfolding demonstration. If we followed a representational view of cognition and Shannon's formal model of a communication system, we might say that signals (or representations) carrying some amount of information about how to use the hand drill were transmitted from the facilitator (the sender or transmitter) to the potential learner (the receiver or destination). Note that this is not unlike saying that, in "production imitation" (**Table 5-1**), the novel action is "acquired by the observer directly through observation" (Hoppitt & Laland, p. 73). However, if we take the transmission metaphor seriously, we should ask, What exactly are these information-carrying signals? How are they transmitted? How does the facilitator encode semantic information about the tool in these signals? How does the potential learner recover the original meaning encoded in these signals? I find these questions, that follow from adopting the transmission metaphor and a representational view of the mind, confusing and unhelpful. My intention here is to provide a positive

account consistent with my (meta)theoretical commitments. Thus, I suggest the following alternative:

1. Successful tool use requires that the person has the relevant abilities of perception-action.
2. Before the activity, only the facilitator, but not the children, has the bodily organization underlying these abilities.
3. During the activity, the facilitator influences the development of these abilities in the children.
4. After the activity, both the facilitator and the children have the bodily organization underlying these abilities.

This way of speaking is consistent with a processual, developmental systems perspective, and with radical embodied cognitive science. The ability to use the tool is seen as depending, not on mental representations ('information' about the task encoded in the brain) but on bodily organization, which cannot be transmitted from one person to another. Children develop their own tool-using ability, i.e. their own bodily organization, as a result of a relational developmental trajectory. The change towards a stronger similarity in how the facilitator and the children used the tool after the activity, compared to before, may justify us attributing them to the same category (e.g. 'successful tool use'), but the metaphor of learning as linear communication of signals makes no sense here.

Characterizing social influences on learning

Making together gave children the opportunity to embody implicit knowledge about using a novel tool and explicit knowledge about the steps required to make a toy vehicle. The results indicated that engaging in the activity indeed transformed children's abilities in both regards. In Chapter 5, I suggested three criteria to clarify the scope of any study of social learning and five criteria to clarify the 'social influences' under scrutiny. The current study provides an opportunity to clarify how these criteria can be used in concrete cases. I first clarify the scope of interest. With regard to the time scale, here I focused on learning that might occur within the short time scale of several minutes. With regard to the spatial scale, the learning situation unfolded within a science museum, in a room enclosed within the exhibition area with no sound insulation, thus providing an environment less controlled than lab-based studies. The main components under analysis were the facilitator, the learners, and the supporting materials. With regard to learning outcome, this study focused on skill or implicit learning (using a novel tool) and explicit learning (verbally identifying the steps in the activity).

I now turn to characterizing the 'social influences' present in this case. With regard to the environmental information made available by the facilitator, this included optic information specifying the tools and materials as well as her actions, and auditory information created by her speech. With regard to the general role played by the facilitator, this must be characterized separately for each learning outcome. With regard to using a novel tool (implicit learning), on the one hand the facilitator acted as a demonstrator and, on the other hand, she provided assistance as the learners attempted to use the tool. Thus, the facilitator engaged both in the same and in a different task than the one in which the learners' performance was judged. With regard to identifying the steps in the activity (explicit learning), the facilitator did not act as a demonstrator but rather as a partner and a guide. Thus, even though the facilitator engaged in the making task together with the learners, their roles were not the same but complementary. Throughout the activity, the facilitator could influence the learner and the learner could also influence the facilitator, thus the influence was bidirectional. The learner interacted with the facilitator in real time, thus the influence between them was direct. The behavioural coupling between them was informational, i.e. by means of perceiving each other, and sometimes also mechanical. With regard to their prior intentions, the facilitator explicitly intended her behaviour to promote learning and thus her behaviour falls under the definition of teaching I put forward in Chapter 5. The potential learners explicitly committed to engage in the activity but there was no commitment to learn something in the process.

Behavioural coordination and learning

This study not only provided analytical tools to examine signatures of coordination but also the association between quantitative measures of coordination and learning. The prediction was that a stronger behavioural coordination would provide a better context for knowledge to be 'passed on'. More specifically, I predicted that children who composed dyads showing higher measures of coordination, compared to lower measures, should improve more. The results contradicted this. To make sense of this result, consider that coordination variables such as RR capture the nonlinear relationship between two time series and characterize the coupled system as a distributed whole rather than the individual systems. In some studies of attentional coordination, participants interacted with recordings rather than live people (e.g., D. C. Richardson & Dale, 2005). In this case,

variation in coordination measures is produced by variation in the participant's behaviour only because it is the only time series that varies in the analysis. In the current study, children interacted with 'the same' facilitator. However, the facilitator did not behave exactly the same with each child because her role explicitly required that she responded to each child differently (**Table 6-2**). Thus, even though she followed 'the same script' every time, she never did things exactly the same way. As a result, the time series representing the facilitator's behaviour is different in every session. Thus, the children and the facilitator contributed to variation in the coordination variables. Compared to dyads with lower values of coordination, a higher value in this case might indicate either (1) that *the child* coordinated his attention to the facilitator more, (2) that *the facilitator* coordinated her attention to the child more, or (3) that *both the child and the facilitator* coordinated more with each other.

Post hoc informal, qualitative analysis of the recordings suggested that when children showed more signs of struggling with the novel tool in the subtask MAKE_WHEELS, the facilitator responded by intervening more and providing more assistance. This was reflected in higher values of RR_{MW} being computed in these cases. Therefore, it might be that the facilitator systematically varied the amount of influence she exerted while assisting children in using the novel tool, depending on the general level of ability demonstrated by the children on the spot. If this was the case, then children whom the facilitator perceived as generally less skilled (as opposed to more skilled) received more attention from her, thus producing higher values of RR_{MW}. Later, in Test-2 these less skilled children took longer to succeed. These considerations suggest that real-time, contingent behavioural adjustments are indeed a central aspect of facilitating-learning relations, at least in some cases. However, studies of social learning and teaching commonly neglect this real-time coordination (Caro & Hauser, 1992; Hoppitt et al., 2008; Hoppitt & Laland, 2013). Thus, further work using similar methods might indeed help elucidate relevant causal processes which are being neglected in the literature. For example, we might ask whether different categories of social learning produce systematically different signatures of behavioural coordination. We might also ask whether signatures of coordination vary across different age groups as they engage in similar learning tasks. For example, suppose we study caregiver-infant dyads playing with a novel tool and that we compare cases with younger and older children. One

hypothesis could be that, in dyads with younger children, the adults play a more prominent role in establishing and maintaining coordination as they interact and that, as they grow older, children might become increasingly able to influence how these situations unfold. These changes might be captured by changes in CRQA variables as well as in diagonal-wise cross recurrence profiles.

Final considerations and link to the following chapter

This study explored new ways to examine behavioural coordination, coupled attention, and social learning within a relatively naturalistic, noisy, environment. One limitation of this design was the lack of control for the explicit knowledge test. Another limitation was the use of relatively broad categories of gaze direction (**Table 6-3**). Consider the category *Learner gazes at Facilitator's eyes*. This category merely indicates that the child is looking in the direction of the adult's eyes with no further qualification, for example, regarding the affective character of the gaze in each case. On the one hand, defining the different categories in such broad terms allowed me to capture dynamical aspects of behavioural coordination and provide evidence of attentional synchronisation throughout the activity. On the other hand, thus defined, this category tells us very little about what each instance of a gaze might mean. A child might look at the instructor's eyes because he is attentive to what the adult is saying, distracted from the activity, excited to be using a hand drill, bored with how the activity is being conducted, and so on. In addition, the current design did not allow me to distinguish between the effects of demonstration and assisted practice on learning. Finally, because both participants contributed to the variation in the coordination variables, this made it difficult to interpret the observed association between coordination variables and learning. The next chapter reports a study which was designed to deal with these limitations. The conditions to be compared were more clearly defined and an eye tracking system was used to produce the gaze direction data, thus avoiding hand-coding from the videos. In addition, potential learners received the same stimulus, thus variation in the coordination variables can be interpreted as indicating variation in the learners and not the facilitator. On the one hand, these changes in the methodology allowed for a more controlled study. This, however, came at the price of losing the messy richness of live interaction.

Chapter 7 Learning through demonstration: a study of eye movements during observational learning of a manual task

7.1 Introduction

In Chapter 6, I focused on learning through making together as a concrete example to advocate for a view of social learning as a relational developmental process that changes embodied skills, rather than as the acquisition or transmission of mental representations. In this chapter, I focus on learning how to solve construction puzzles through observing demonstration videos as another concrete example. Similar to producing a toy vehicle, solving a construction puzzle can be regarded as a making activity. Here, too, the maker manipulates a set of materials, thus producing forces that change the spatial arrangement among them and bring about a change in form. And here, too, the bodily organization that underlies the abilities that the maker brings to the activity persist under transformations, possibly changing in systematic ways that we observers might wish to describe as learning. However, solving a construction puzzle is different from making a toy vehicle in that the form of the final product is much more constrained. In the previous study, participants were invited to produce their own toy vehicle and decorate it however they liked (within given constraints). Here, participants were expected to produce the pre-defined forms shown to them, which would be considered the correct solution to the puzzles.

Watching an expert (henceforth, ‘facilitator’) engage in some task is one among many possible ways in which skill learning can be influenced by the presence and activity of others around us. Learning by observing others can occur in formal settings such as in schooling, more informally such as when we watch instructional videos online showing

us how to prepare a recipe, or during our encounters with those around us as we go about our daily activities. In the dominant literature on social learning, it is common to read that, in such situations, facilitators transmit 'information' (*sensu* represented knowledge) about the target task to learners or that learners acquire such representations socially rather than asocially (e.g., Csibra & Gergely, 2009; Hoppitt et al., 2008; Hoppitt & Laland, 2013). However, from the perspective put forward in this thesis, the knowledge that experts 'have' is embodied in their physical organization and therefore cannot be 'transmitted' except in a very loose, metaphorical sense. Rather, learners must construct their own knowledge themselves as they encounter the world.

One of the main goals of this chapter is to further clarify why the transmission metaphor is inconsistent with an approach committed to developmental systems thinking and radical embodied cognitive science, and to offer an alternative, positive account. How, then, one might ask, can the facilitators affect the learners if not by 'transmitting information' to them? The alternative being put forward relies on the notion of ecological information (J. J. Gibson, 1979/2015) rather than representations, and views ontogeny as a history of causal relations among the components of the organism-environment system in which some abilities persist and others change. The basic idea is that, by their presence and behaviour, facilitators can structure the ambient arrays of light, sound, chemicals, objects, and so on, and therefore create environmental information (*sensu* meaningful patterns, not 'knowledge') that is publicly available (J. J. Gibson, 1979/2015). Learners may or may not detect these patterns and use them to guide their own actions. Depending on how they make sense of what they perceive in the changing situations in which they find themselves, learners may exercise their existing abilities in new ways, and even possibly develop new ones, thus embodying new knowledge. Learning, then, can be viewed as a historical, relational process of "education of attention" (J. J. Gibson, 1979/2015) rather than a computational process of accumulation of representations. It is a category of developmental process, one in which the bodily structures that underlie the individual's abilities (bones, joints, muscles, neuronal networks, and so on) change and become attuned to their environment in different, maybe even relatively improved, ways.

The present study focused on two aspects of the environmental information made available by the facilitator that potential learners might pick up and use to guide their own

behaviour. The first aspect is the information (*sensu* meaningful patterns in the light) made available in the optic flow specifying the facilitator's face. The second aspect is the information (*sensu* meaningful patterns in the air waves) made available in the sounds produced by the facilitator's verbalizations. These aspects are related to the mechanisms underlying joint attention, a term denoting those cases in which the individuals are aware that they are attending to the same unfolding situation. Joint attention is widely recognized as fundamental to social relations, including early language acquisition (M. Carpenter, Nagell, & Tomasello, 1998; Tomasello, 2003), verbal communication (Clark & Krych, 2004; Dale, Kirkham, & Richardson, 2011; D. C. Richardson & Dale, 2005), and joint action (Kraut, Fussell, & Siegel, 2003; Sebanz & Knoblich, 2009). Gaze following is considered central to establishing and sustaining joint attention (M. Carpenter et al., 1998; Tomasello et al., 2005). This is the case because, by looking at other people's face and eyes, one might detect the invariants specifying where *they* are looking and thus be informed about what *they* are looking at. Depending on one's intentions, one can use this ecological information to adjust one's own behaviour and attention accordingly.

However, in goal-directed actions, such as reaching and object manipulation, eye movements and hand movements are tightly correlated in space and time. People tend to look at what they are manipulating (Johansson, Westling, Backstrom, & Flanagan, 2001; Land & Hayhoe, 2001). Additionally, people tend to shift their visual attention to the target object or location just before initiating reaching, and they tend to continue looking at it until the movement is completed (Horstmann & Hoffmann, 2005; Land & Hayhoe, 2001). In contexts such as these, both the eye movements and the hand movements are linked to the actor's current target of attention by constraints and can therefore inform about it. Thus, when we look at people manipulating objects, we may not need to examine their eyes to learn about where they are looking – we can just follow their hands instead.

This mechanism leading to joint attention is not restricted to, say, adults with extensive experience in observing manipulative actions. Yu and Smith (2013) provided evidence for this route to joint attention in the case of one-year-old children and their parents as they played with toys while both were being eye tracked. In this context, the participants rarely looked at each other's faces and, nevertheless, were able to establish

and sustain joint attention by simply looking at their own hands (when manipulating objects) or at the partner's hands (when the partner manipulated objects). Thus, this process is possibly much more widespread than previously acknowledged.

Another aspect of how individuals may coordinate their attention and behaviour is related to the role of speech. One of the primary roles of speech is to influence the attentional, and therefore also the intentional, state of the listeners (Briggs, 2002; Tomasello, 2003). During a conversation, speakers and listeners monitor their own actions (which include their vocalizations) as well as the actions (including the vocalizations) of the conversational partner, adjusting their behavioural flow as they go along (Clark & Krych, 2004; Fogel, 1993). Evidence shows that speakers tend to look at those objects that correspond with the words being spoken (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998), and that listeners also tend to look at those objects that correspond with the words being heard (Allopenna, Magnuson, & Tanenhaus, 1998; D. C. Richardson & Dale, 2005). As a result of this mutual co-regulation of behaviour, the attention of speakers and listeners can become more synchronized over time (D. C. Richardson et al., 2007). Thus, both the possibility to see the partner's face, and the possibility to listen to the partner's vocalizations, may impact how people detect and use the information (*sensu* meaningful patterns) available in their environment to guide their own behaviour, therefore possibly influencing how they learn about the target task.

Previous work on coordination of attention and task performance

A few studies have used eye tracking explicitly to address the relation between coordination of visual and/or auditory attention, and performance in some task. A study by van Gog, Jarodzka, Scheiter, Gerjets, and Paas (2009) used a computer-based problem-solving task known as 'Frog Leap' and examined whether attempting to direct the learner's attention during a learning phase improved performance in a subsequent test phase. To direct the learners' attention, the authors first asked a skilled demonstrator to solve the problem didactically while using an eye tracker, and simultaneously captured (1) the computer screen showing the task being solved, (2) the audio of the demonstrator's verbal explanation of his actions, and (3) the eye movements (i.e. fixations on the screen) of the demonstrator. Following a 2 x 2 factorial design, the authors then produced four types of example videos to be used as stimuli: (1) screen capture only; (2) screen and

audio capture; (3) screen capture and demonstrator's eye movements; and (4) screen and audio capture and demonstrator's eye movements. The eye movements of the demonstrator were shown as a coloured dot moving on the screen, and the authors assumed that this would guide the learners' attention. Participants were exposed to one of the four conditions. First, they went through a learning phase where they received the stimulus twice. They were then asked to solve the problem twice in the test phase.

The authors predicted that showing the facilitator's eye movements would contribute to learning, and that this effect would be larger in the group receiving visual and auditory stimulus compared to visual stimulus only. Neither of these predictions were observed. Results suggested that showing the demonstrator's eye movements superimposed on the screen required higher investment of mental effort by the learners and, contrary to expected, was either neutral or detrimental to learning. In addition, regardless of attention guidance, no difference in performance was seen between participants receiving visual stimulus only and participants receiving both visual and auditory stimulus. However, the authors did not provide further interpretation of this finding.

D. C. Richardson and Dale (2005) reported two studies examining the coupling of visual attention between speakers and listeners. In the first study they recorded the speech and eye movements of four participants while they looked at a set of pictures of famous cast members of a TV sitcom (either ''Friends'' or ''The Simpsons'') and spoke spontaneously about them. They then played one-minute segments of the recorded speech to a separate set of participants who were looking at the same pictures and also recorded their eye movements. Results from cross-recurrence quantification analysis showed that speakers and listeners synchronized the direction of their gaze above chance with a peak at a 2-sec lag, consistent with the speaker leading the listener's attention. In addition, a stronger coupling was associated with higher scores in a comprehension test subsequently applied to the listeners. In the second study reported, the authors attempted to direct the listener's attention by flashing (i.e. changing from dimmed to full colour) the pictures of the characters, either at the time that the speaker had looked at them or, in the control condition, following a randomized order. Although no difference was seen in the overall accuracy of listeners' comprehension between conditions, listeners in the synchronized

condition were quicker in answering the comprehension questions. The results from this systematic manipulation provide some evidence that the attentional coordination between speakers and listeners is not only correlated with, but plays a causal role in, discourse comprehension. In other words, listeners tend to comprehend better what a speaker says when they look at the same objects at the same time.

Grant and Spivey (2003) used an insight problem known as Duncker's radiation problem to investigate whether directing the attention of participants in a systematic way can facilitate learning. The solution to insight problems cannot be logically induced and is often arrived at suddenly by the learners, who often describe an accompanying 'Aha!' sensation. In the first study reported, the authors recorded the eye movements of participants as they attempted to solve the problem. Results showed that, compared to unsuccessful participants, successful problem solvers spent more time looking at the same particular area of the task diagram in the 30 s before arriving at the solution. In the second study the authors attempted to direct the attention of participants by subtly blinking and therefore increasing the perceptual salience of the critical area of the task diagram. As predicted, compared to highlighting a non-critical area and to presenting a static stimulus, highlighting the critical area facilitated insight and led to increased success rates.

Recently, Coco, Dale, and Keller (2018) investigated the association of gaze coordination in dyads and performance in a spot-the-difference task. Participants viewed a set of scenes on their individual monitor and, on each trial, had to decide whether they were viewing the same or different scenes. In one condition they could talk freely to reach a consensus. In another condition, only one partner could speak while the other could give no feedback and had to remain silent. In a third condition, the listener was allowed to give only minimal feedback in the form of backchannel to signal understanding. The eye movements of participants were recorded and their coordination examined using cross recurrence quantification analysis. The authors provide evidence of an association between coordination of eye movements and performance, but only in the condition in which the listener could not talk to the speaker. Specifically, two of the measures of coordination they examined were negatively associated with performance. The authors interpreted this negative association by arguing that, as the nature of the task was akin to a visual search task, if the eye movements of the participants are too much aligned this

suggests that the listener is following the speaker's verbalisation too closely and not exploring parts of the images that are potentially relevant.

In sum, the evidence of an association between attentional coordination and task performance is scarce and mixed. D. C. Richardson and Dale (2005) provided some evidence that attentional coordination was positively associated with discourse comprehension. In contrast, Coco et al. (2018) provided evidence of a negative effect that nevertheless seemed consistent with the context of their task. Similarly, studies attempting to direct the learners' attention by means of the stimuli also showed mixed results. By highlighting the critical part of the stimulus, Grant and Spivey (2003) managed to increase success rate. In contrast, in the study by van Gog et al. (2009), showing a coloured dot on the screen to indicate where the facilitator was looking at was, if anything, detrimental to learning. However, in this study the authors did not analyse the eye movements of the learners and therefore presented no evidence to support the premise behind the methodology, namely that the attention of the learners was indeed guided by the coloured dot on the screen. It is possible, as the authors admit, that the extra visual stimulus superimposed on the demonstration video might have distracted the learners rather than guided their attention.

There is no reason to expect a simple and uniform picture of how attentional coordination affects learning because both attentional coordination and learning can refer to many related processes. This class of phenomena is likely to be influence by many factors including the nature of the task being examined, the abilities of learners and facilitators, the affordances available in each situation and specified by the patterns in the ambient arrays, how the individuals might affect each other in each case, what counts as coordination, and so on. The present study aims to contribute to this literature and, more broadly, to the growing field of radical embodied cognitive science (Chemero, 2009; J. J. Gibson, 1979/2015; Thompson & Varela, 2001; Varela et al., 1991).

7.1.1 The present study

Here I examine the dynamics of eye movements during observation of demonstrations of the solution to a novel manual task in adult humans, and its association with learning. Participants watched demonstration videos and then attempted to solve a target task. This

procedure was repeated five times with each task, and this iterative procedure was repeated with three different tasks. I used eye tracking and cross recurrence quantification analysis to examine dynamical features of the learners' eye movements as they watched the demonstration videos, and I examined their subsequent performance at the task in each trial iteration. I investigated whether seeing the facilitator's face and listening to his vocalizations affected (1) how learners coordinated their attention to the actions being demonstrated and to the facilitator's face, and (2) how they performed at the task. I also investigated whether measures of attentional coordination were associated with learning, here operationalised as improved performance with practice.

Predictions

I have indicated two, non-mutually exclusive, mechanisms associated with joint attention above: gaze following (e.g., M. Carpenter et al., 1998) and hand-eye coordination (Yu & Smith, 2013). To test which might be important here, I examined whether being able to see the facilitator's face, compared to not (face blurred), affected how learners coordinated their visual attention with the actions being demonstrated. If gaze following was important for establishing and sustaining attentional coordination, I expected higher values of coordination in the conditions where the facilitator's face was visible compared to blurred. However, if gaze following was not important for attentional coordination in this task, I expected no evidence of differences in coordination measures between these conditions.

I also examined whether listening to the facilitator, compared to not, affected how learners coordinated their visual attention with the actions he was demonstrating. The explicit communicative intention of the facilitator in this case was to describe the actions required to solve the task as they were performed (see Appendix). I therefore predicted that learners who could listen to the facilitator, compared to not, would show higher values of coordination.

Finally, I examined the association between coordination of attention and learning. Given the nature of the task, looking at precisely those pieces being manipulated should allow learners to detect relevant aspects of the demonstration, and this should be beneficial to learning. I therefore predicted that, overall, higher values of coordination

between the pieces being manipulated by the facilitator and the pieces being looked at by the learner would be beneficial to learning.

The facilitator's behaviour during the recording of the demonstration videos was certainly influenced by his own prior experience in similar situations, including his interaction with schoolchildren with whom he has previously engaged as a teacher. However, he did not interact with the potential learners in this study directly. Therefore, there is an important asymmetry in this design: while the behaviour of the facilitator might influence the behaviour of the learners, the opposite cannot occur. Thus, the design does not examine bidirectional behavioural coupling but unidirectional coupling in which the direction of influence is from facilitator to learners. It is in this sense that coordination and social influence should be understood. The benefit of this design is that, because all participants receive the same stimuli, the observed variation is restricted to differences among the learners. The weakness that comes with it is that the constrained, lab-based learning situation thus created might evoke behaviours different from a more naturalistic situation. Therefore, as in any empirical study, generalizations should be made with care.

7.2 Method

7.2.1 Site and participants

The study took place at the Joint Eye-tracking lab at the University of Edinburgh (UK). An experienced school teacher in Edinburgh (33 years of age, male) was recruited to perform the role of the facilitator in the video recordings used as the stimuli and received £20 for his time. A further 53 adult participants were recruited to perform the role of learners using the Experimenter Volunteer Panel of the University of Edinburgh and by emails sent to students and staff in different departments. Ages ranged from 18 to 29 with two outliers of ages 42 and 50 (median: 21; s.d. = 5.4), and 32 (60%) were female. All participants gave informed consent, had normal or corrected-to-normal vision, indicated no known learning disability, and were paid £7 as compensation for their time. Each experimental session lasted for 45-60 minutes. The study was approved by the University of St Andrews Teaching and Research Ethics Committee and by the Psychology Research Ethics Committee of the University of Edinburgh, in accordance with the British Psychological Society guidelines on ethics.

7.2.2 Stimuli

The task consisted of construction puzzles, which are solved by assembling a set of pieces to form specific structures. Three different commercially available wooden puzzles were used, which differed in the number of pieces (star puzzle: six pieces; egg puzzle: eight pieces; barrel puzzle: twelve pieces) and in the sequence of steps required to solve them. The stimuli used to potentially facilitate learning were demonstration videos and, in the case of the control conditions, neutral still images extracted from the initial frames of these videos (i.e., prior to beginning the demonstration; see **Figure 7-1**). The participant acting as the facilitator was shown how to solve each puzzle by the experimenter and was given time to practice. To produce the demonstration videos, the facilitator performed explicit demonstrations of each of the solutions, accompanied by verbal instructions of the movements being performed. The script of each demonstration was specified beforehand by the experimenter and the facilitator together in order to achieve a naturalistic sequence and to ensure the demonstrations were similar in structure and vocabulary across the three tasks (transcriptions of the verbal instructions are available in the Appendix.

A tripod-mounted camera was positioned at eye level in front of the facilitator, where an imaginary learner would stand. The facilitator was instructed to act naturally and to look at the camera from time to time, as if he were teaching an imaginary learner in front of him. The videos were captured in the portrait orientation and a lapel microphone was used to capture the facilitator's speech. Several shots were recorded for each puzzle and the most natural and fluent one was chosen to be further processed. Because the puzzles differed in the number of pieces, the final videos differed in duration (star: 40s; egg: 54s; barrel: 78s).

To produce the stimuli for the four experimental conditions, the initial demonstration videos were edited in the Wondershare Filmora software. For each puzzle, four different videos were produced. In two of them, the face of the facilitator was blurred while no image editing was performed in the other two. Furthermore, in two of them the sound was completely muted while no sound editing was performed in the other two. This yielded four versions of the demonstration video for each task, corresponding to the four experimental conditions described below. Furthermore, to produce the stimuli for the two

control conditions, another two videos were produced for each task consisting of a single still image extracted from the initial frames showing the facilitator in the neutral starting position (with his hands on the desk without touching the puzzle pieces) and which either included or did not include the corresponding sound track.

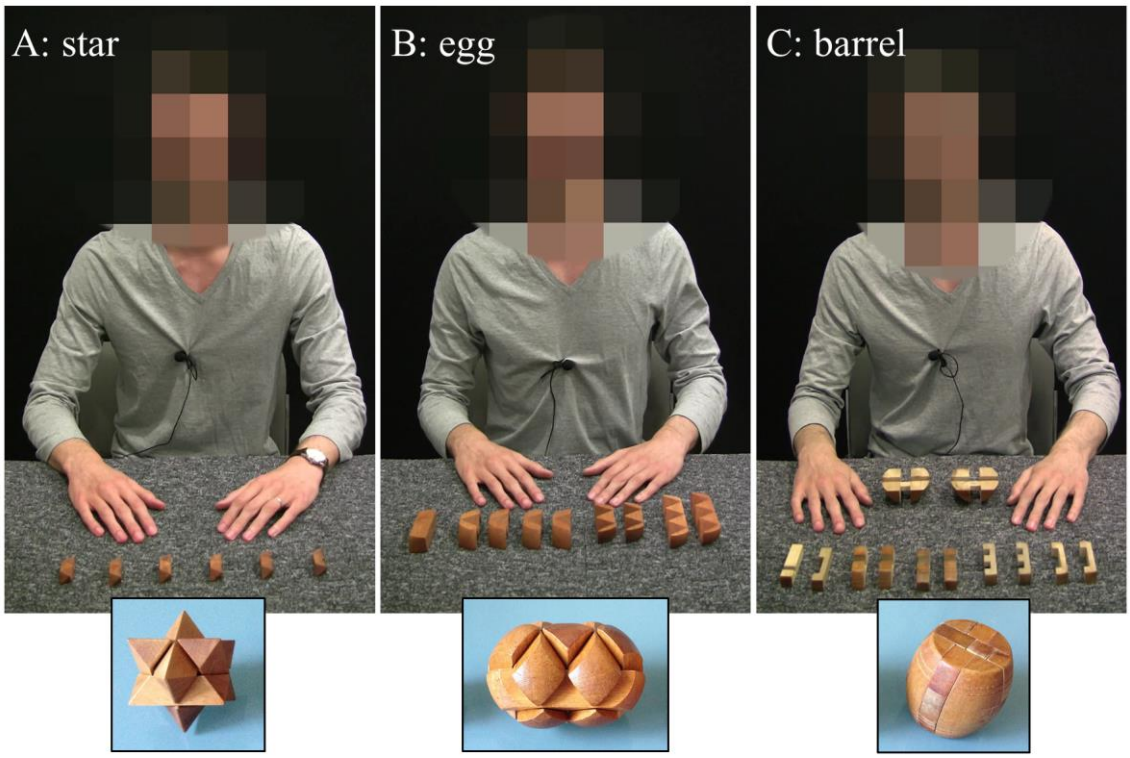


Figure 7-1. The three tasks (A: star, B: egg, C: barrel) and samples from the stimuli (demonstration videos) used in the study. The still images were extracted from the initial frames of the demonstration videos and show the facilitator in the neutral starting position, i.e. with his hands on the desk without touching the puzzle pieces in front of him. The insets show the corresponding solved puzzles.

7.2.3 Design

The experimental conditions varied in a 2 (facilitator's face visible, face blurred) by 2 (with audio, no audio) design. This produced four conditions: face blurred and no audio (noFACE_noAUDIO); face visible and no audio (FACE_noAUDIO); face blurred and with audio (noFACE_AUDIO); and face visible and with audio (FACE_AUDIO). In addition, I ran two extra conditions in which participants were shown a neutral still image of the facilitator and the puzzle pieces rather than a video (similar to the snapshot shown bottom left in **Figure 7-3**).These conditions varied in terms of the audio that accompanied it, thus producing the conditions no video and no audio (noVIDEO_noAUDIO); and no

video and with audio (noVIDEO_AUDIO). Participants were randomly allocated to one of the six conditions.

All participants faced the three versions of the task (star, egg, and barrel puzzles) and the order of presentation was counterbalanced across participants. Participants engaged in each task in five consecutive iterations (after an initial baseline test), each comprising presentation of the demonstration video followed by the opportunity to try solving the task. This iterative procedure is interesting for the study of learning because it produces repeated measures forming a learning curve rather than just a one-off success/failure outcome.

7.2.4 Experimental setup

Participants manipulated the puzzles on one desk and were exposed to the stimuli at another desk, as shown in **Figure 7-2**. Participants could easily move between the two desks by simply rotating their chair 90 degrees and adjusting for comfort, ensuring that they could alternate between engaging in the tasks and watching the demonstration videos with little disturbance. Videos were displayed on a 21'' monitor in portrait orientation at a resolution of 1680 x 1050 pixels at a refresh rate of 100 Hz and a frame rate of 25 Hz, and the audio media were played on standard desktop speakers.

One eye tracker (SR Research EyeLink 1000 with Desktop Mount) was used with a sampling rate of 1000Hz. A chin and headrest were used to limit head movement. The monitor covered 35 degrees of visual angle vertically and 22 degrees horizontally, and the distance between the headrest and the top of the monitor was 74 cm. The experiment was implemented in the SR Research Experiment Builder software. All sessions were also video recorded using two tripod-mounted cameras, but these images were used only to double check measures of success initially coded in real time during the sessions.

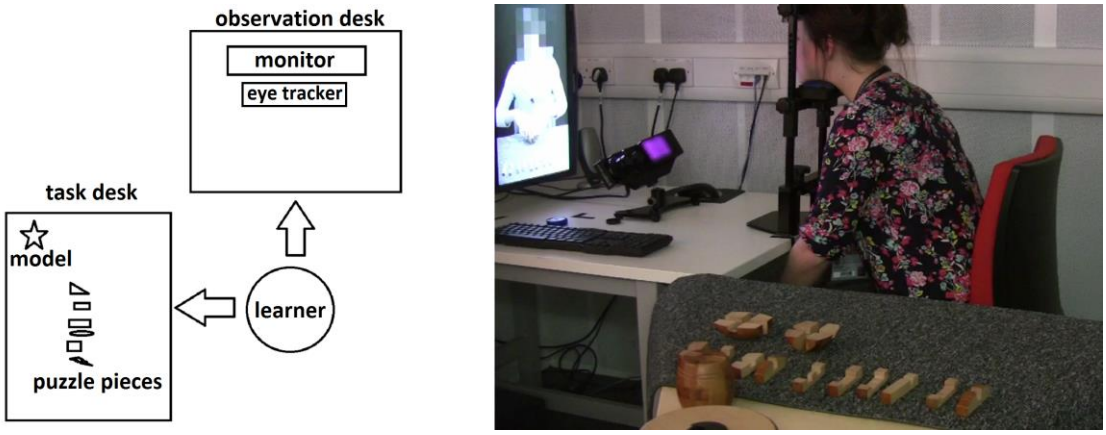


Figure 7-2. The workspace. The learner is at the observation desk watching the demonstration video while being eye tracked. To her left is the task desk with a solved barrel puzzle (model) and the corresponding pieces. The other two puzzles (star and egg) are currently hidden from the participant's view.

7.2.5 Procedure

Participants were informed that they would alternate five times between trying to solve a puzzle and attending to the visual and/or auditory stimuli while their eye movements were recorded, and that this procedure would be repeated for three different puzzles. The experimenter placed the pieces of the first puzzle in front of the participant, as well as an assembled model, and asked if the participant had seen and/solved it before. The participant was then given the instruction to attempt to solve the puzzle as quickly as possible, i.e. to manipulate the pieces such as to produce a copy of the assembled model. They could inspect the assembled model visually but were not allowed to touch it. Learners were given a fixed time interval to solve the task (star: 90s, egg: 90s, barrel: 120s), corresponding roughly to twice the time required by the facilitator to solve them.

The first attempt was performed before any stimulus was presented and was therefore used as the baseline score. After this, the participant turned to the observation desk, positioned him or herself on the head and chin rest and the appropriate stimulus (video or still image, with or without face, with or without sound) was then presented. During this period, the experimenter disassembled the partially or completely solved puzzle and arranged the pieces randomly on the desk for the next attempt. The participant then turned back to the task desk with the same instruction to attempt to solve the puzzle again as quickly as possible. This procedure of alternating between attending to the stimulus and attempting to solve the task was iterated five times in total before the same procedure was performed for the second (and then third) puzzle.

At the beginning of the session a parallax test was conducted and only the data for the dominant eye were used in the analysis. Nine-point calibration routines were performed before the first attempt at each puzzle, and a drift correction was performed before each iteration. The experimenter coded the participant's performance at the end of each attempt as either a success (i.e. puzzle was assembled correctly within the time given) or no success (i.e. task was not solved).

7.2.6 Data processing

Facilitator's manipulation and gaze data

I coded the facilitator's behavioural flow from the demonstration videos into categorical time series at a sample rate of one observation every 25 ms using the free software Solomon version beta 17.03.22 (Péter, 2016). For each puzzle solving demonstration, I produced two time series. The first included labels for the individual puzzle pieces and for the aggregates (partially solved puzzle) that appear as the pieces are being assembled. It also included a label to indicate no manipulation. An example is shown in **Figure 7-4**, in which different colours indicate the different labels. Often the facilitator held the puzzle in one hand while reaching for the next piece to be incorporated. In these cases, the label for the new piece was used. This time series was used together with the corresponding time series indicating the pieces or aggregates being fixated by the learners.

The second time series produced for each demonstration video included one label for when the demonstrator looked directly at the camera (thus simulating the situation of looking at the learner) and one label for 'other'. This was used together with the corresponding time series indicating when the learners looked at the facilitator's face (see **Figure 7-6** for an example). The length of these time series was normalized according to the length of each of the videos (star: 1500 bins, egg: 2000 bins, barrel: 3000 bins).

Learner's performance data

I coded the binary performance data (1 for 'success' or 0 for 'no success') during the trials and later confirmed these data with the video recordings of the sessions.

Learner's gaze data

Only the gaze data of participants in conditions noFACE_noAUDIO, FACE_noAUDIO, noFACE_AUDIO, and FACE_AUDIO (i.e. the 2 x 2 = 4 experimental conditions under

scrutiny) were used in the analyses reported below. Fixation events were extracted with the SR Research parsing algorithm using the Data Viewer software at its default parameter settings. For each video I defined regions-of-interest, including a margin of one degree, around each puzzle piece, the emerging aggregates (partially solved pieces), and the facilitator's face. The raw gaze recording yields one observation every millisecond. To reduce this unnecessarily high resolution, I produced scan patterns with one observation every 25ms and then normalized them to a fixed length to avoid small differences that occur across trials. Scan patterns are categorical time series representing the sequence of fixated targets. Because the demonstration videos for each puzzle varied in length, so did the final time series. Their lengths were: star: 1500 bins; egg: 2000 bins; barrel: 3000 bins. Each bin contributes one value in the final time series, which is a label of the target being fixated by the learner most of the time in that brief interval of about 25ms.

Similar to the facilitator's data, I also produced two time series to represent the learner's gaze in each iteration. The first included labels for the individual puzzle pieces and for the aggregates (partially solved puzzle) that appear as the pieces are being assembled. It also included a label for 'other'. This time series was used to examine how participants coordinated their visual attention to the facilitator's actions. The second time series included a label for when the learner looked directly at the facilitator's face, and one label for 'other'. This time series was used to examine how participants looked at the facilitator's face. I also computed the proportion of fixation time that learners spent looking at the facilitator's face and at the puzzle pieces. **Figure 7-3** illustrates the data from which the time series were produced.

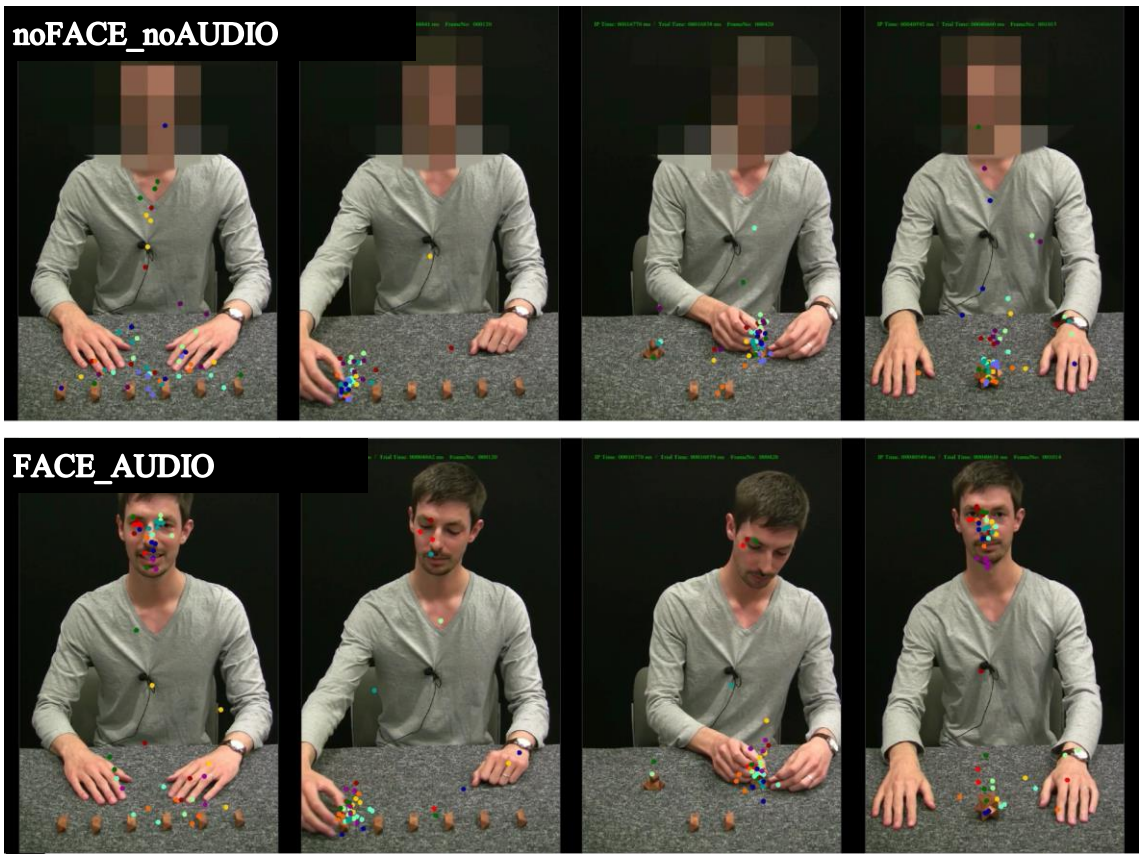


Figure 7-3. Snapshots from the demonstration of the star puzzle in two conditions (noFACE_noAUDIO and FACE_AUDIO). The coloured dots represent the eye tracking data from all participants in each condition. The cloud of dots indicate that participants tended to be looking at the pieces being manipulated by the facilitator. It also indicates, especially in the first and last frames, that participants looked more at the facilitator's face when it was clearly visible.

7.2.7 Analysis

I examined, from each trial, the performance of the potential learners (success or no success); the proportion of fixation time directed at the facilitator's face and at the puzzle pieces; the time series representing the sequence in which the facilitator manipulated the puzzle pieces in the demonstration videos; the time series representing when the facilitator looked at the camera; the time series representing the sequence in which the learners looked at the puzzle pieces while watching the demonstration videos; and the time series representing when learners looked at the facilitator's face in the videos.

The analysis of difference in performance across all six conditions did not involve eye tracking data. In this case, from the initial 795 trials (53 participants x 3 puzzles x 5 iterations), only 5 trials of a single participant were excluded because she was familiar with the star puzzle.

For the remaining analyses involving eye tracking data, however, we focus only on four experimental conditions and therefore start with an initial dataset of 600 trials (40 participants x 3 puzzles x 5 iterations). From these, 3 trials were excluded due to early termination (the participant moved away from the eye tracker), 5 due to the participant knowing the puzzle (same case as reported above), 124 due to the eye tracking data not being acquired properly, as indicated by more than 20% of out-of-range fixations (103 cases) or by artefacts appearing in the time series (21 cases). A further two cases were excluded due to the participant inadvertently moving the desk during data collection. The final dataset included in the results reported below included 466 trials. The number of participants was well balanced across conditions: noVIDEO_noAUDIO = 6; noVIDEO_AUDIO = 7; noFACE_noAUDIO = 10; FACE_noAUDIO = 8; noFACE_AUDIO = 8; and FACE_AUDIO = 10.

Recurrence Quantification Analysis

4) Cross recurrence plots and joint recurrence plots

To examine the coordination dynamics between the learners' direction of gaze and the facilitator's behaviour (manipulating the pieces and looking at the camera), I used a family of nonlinear techniques from recurrence analysis based on cross recurrence plots (**CRPs**) and joint recurrence plots (**JRPs**), already introduced in Chapter 6 (Marwan & Kurths, 2002; Marwan et al., 2007; Shockley et al., 2002; Webber & Zbilut, 2005; Zbilut et al., 1998). These analyses were conducted in `R` with the `crqa` package (Coco & Dale, 2014). I have examined coordination in two cases. The first case is the coordination between the sequence of puzzle pieces manipulated by the facilitator and the sequence of pieces fixated by the learner as they watch the demonstration videos. The second case is the coordination between the facilitator looking at the camera and the learner looking at the facilitator's face.

I first focus on the coordination between the facilitator manipulating the pieces and the learners looking at the pieces. For each trial I produced a cross recurrence plot from the two corresponding time series (facilitator manipulating pieces and learners gazing at them) and I extracted quantification variables (similar to chapter 6). Here only the labels indicating the different puzzle pieces, but not the labels for 'other' or missing data, were

allowed to contribute towards cross recurrence. It might be helpful to consult **Figure 7-4**, **Figure 7-5**, **Figure 7-6**, and **Figure 7-7** while reading the next section.

In this study, each cross recurrence point (i.e., each coloured point in the CRP) corresponds to approximately 25ms of the behavioural flows being compared. I examined three coordination variables: recurrence rate (**RR**), determinism (**DET**), and mean line length (**L**). **RR** is the proportion of cross recurrences that obtain in the cross recurrence matrix represented in the CRP. In **Figure 7-5**, the RR corresponds to the number of coloured points divided by the total number of points (i.e., coloured or white) in each plot. This is a general and indiscriminate measure of the degree of shared activity or coordination between the two time series being compared, corresponding to the more widely known cross correlation sum (Kantz, 1994).

DET is defined as the proportion of cross recurrence points that form continuous diagonal lines (sequence of points with slope = 1) longer than some threshold. These diagonal lines are important because they indicate when the systems exhibit a common trajectory in the phase space, i.e., when they visit (roughly) the same state. If the two systems visit the same state at the same time, this is indicated by cross recurrence points forming the main diagonal passing through the origin. Alternatively, if the two systems visit the same state at different times (i.e. with some time lag), this is indicated by cross recurrence points forming diagonals parallel to the main (not passing through the origin). DET is therefore an index of synchronization between the systems. Here, DET quantifies the synchronization between the learner's direction of gaze and the specific piece being manipulated by the facilitator during the demonstration. In **Figure 7-5**, the DET corresponds to the number of coloured points that form continuous diagonal lines divided by the total number of coloured points is each plot. In the present case, higher values of DET indicate that learners consistently gaze more often at the specific piece being manipulated by the facilitator, with or without some lag.

L is defined as the average length of the diagonal lines (longer than the threshold) formed by cross recurrence points. This is an index of the proportion of time in which both systems stay synchronized. In **Figure 7-5**, the L corresponds to the average length of the diagonal lines that obtain in each plot. Higher values of L indicate that, when the learner synchronizes her or his gaze with the behaviour of the facilitator, she or he stays

synchronized for a relatively longer time. Here, each unit of L corresponds to approximately 25ms of synchronized attention.

The values of variables DET and L depend on the threshold parameter `mindiagline`. The appropriate value for this parameter was defined empirically by (1) running the `crqa` function with a range of possible `mindiagline` values, (2) plotting the resulting DET values as a function of the different `mindiagline` values, (3) visually inspecting these plots and (4) choosing the parameter value that reduces the ceiling effect in the DET variable. Following this procedure suggested a threshold value of 30 points was appropriate, which corresponds to about 750 ms in the raw data. In other words, only synchronized attention sustained for longer than 750 ms counted towards values of DET and L. CRQA variables are descriptive in nature and therefore comparisons among conditions, participants, or against appropriate baselines are required to draw inferences and examine specific predictions (Marwan et al., 2007; Shockley et al., 2002).

To capture the iterative character of the experimental design and examine how coordination dynamics evolved across trials, in addition to the cross recurrence plots I also examined joint recurrence plots (JRPs) for every participant and for every task. JRPs were produced, only for iterations 2 to 5, by multiplying the cross recurrence matrices computed for each iteration and all previous iterations with the same task. For the first iteration I simply used the cross recurrence plot (since there is no previous CRP to multiply it with). For iteration 2 I multiplied the two cross recurrence matrices obtained for iterations 1 and 2. For iteration 3 I multiplied the three cross recurrence matrices obtained for iterations 1, 2, and 3; and so on.

Cross recurrence matrices only have values of one and zero. Therefore, multiplying them produces a matrix which also only has values of one and zero. Specifically, if all cross recurrence matrices being multiplied have a value of one in some entry $[x_i, y_i]$, then the resulting joint recurrence matrix will also have a value of one in the corresponding entry. If, on the other hand, there is at least one cross recurrence matrix in which there is a value of zero in that position, then the resulting joint recurrence matrix will have a value of zero in the corresponding entry. Thus, the resulting JRPs indicate *the patterns of coordination dynamics that consistently appear across trials* with each task. I computed

the quantification variables RR, DET, and L, from these JRPs and interpreted them as measures of repeated or consistent attentional coordination across trials.

I used the same procedure just described to examine the coordination between the facilitator looking at the camera and the learners looking at the facilitator's face. For each trial I produced a cross recurrence plot from the two corresponding time series and I extracted the same quantification variables: RR, DET and L. Here only the labels for *facilitator looking at the camera* and *learner looking at the facilitator's face*, but not the labels for 'other' or missing data, were allowed to contribute towards cross recurrence. I also produced the joint recurrence plots as described.

As an illustrative example, I show data from one participant in the condition FACE_AUDIO. **Figure 7-4** shows the time series of the pieces of the star puzzle being manipulated by the facilitator and the time series of the learner looking at the pieces while watching the demonstration video in each trial. **Figure 7-5** shows the corresponding CRPs produced from pairing the time series of the facilitator with each of the time series of the learner, and the corresponding JRPs produced by multiplying the CRPs as described above. In addition, **Figure 7-6** shows the time series of the facilitator looking at the camera as he manipulated the pieces of the puzzle, and five time series of the learner looking at the facilitator's face while watching the demonstration videos in each trial. **Figure 7-7** shows the corresponding CRPs and the JRP produced by multiplying them.

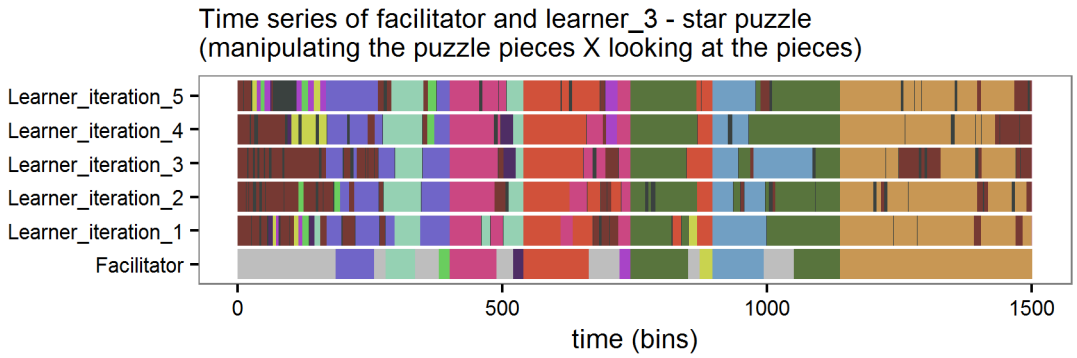


Figure 7-4. Time series of the facilitator manipulating the pieces of the star puzzle and five time series of a learner looking at the pieces while watching the demonstration videos. The colours indicate either a single piece or the partially assembled puzzle being manipulated (or looked at). The grey periods in the facilitator's time series indicate moments in which he was not manipulating any piece. Note that the learner tended to be looking at exactly the piece (or composite) being manipulated by the facilitator at any moment.

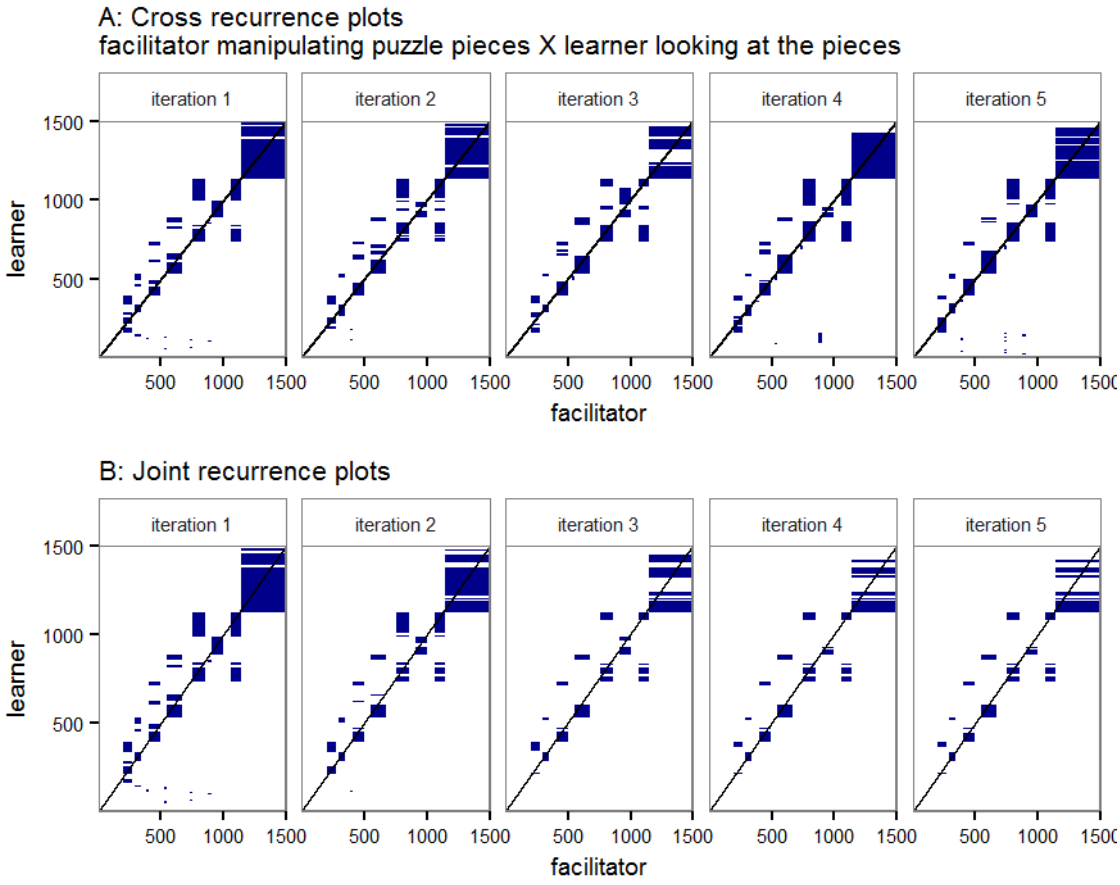


Figure 7-5. A: Cross recurrence plots (CRP) produced from the time series of the facilitator manipulating the puzzle pieces (horizontal axis) paired with the time series of learner_3 looking at the pieces while watching the demonstration videos (vertical axis). The line of synchrony, in which the time stamp of the two time series coincide, is shown in black, and cross recurrences are shown in blue. These are the same cases shown in **Figure 7-4**. B: Joint recurrence plots (JRP) produced from the CRPs shown in A. For each iteration, the JRP is produced by multiplying the CRP from that iteration with all previous ones, thus leaving only those recurrences that occur in all of them. For iteration 1, the CRP is used.

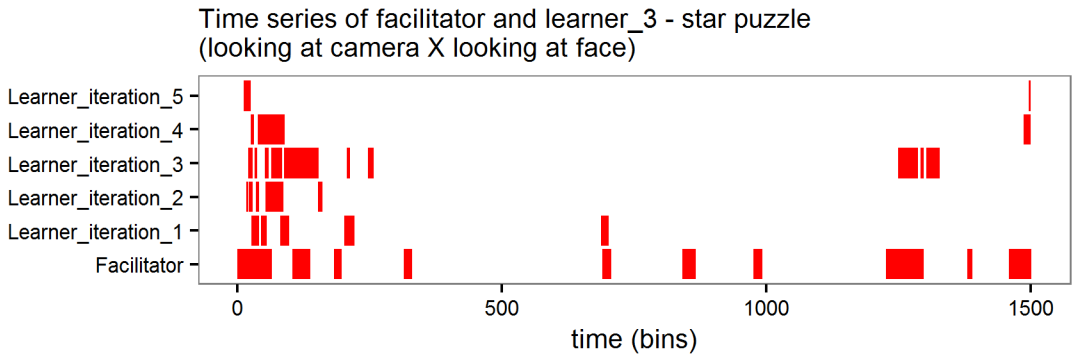


Figure 7-6. Time series of the facilitator looking at the camera as he manipulated the pieces of the star puzzle, and five time series of learner_3 looking at the facilitator's face while watching the demonstration videos. The colour red indicates the relevant states. These are the same trials shown in **Figure 7-4**.

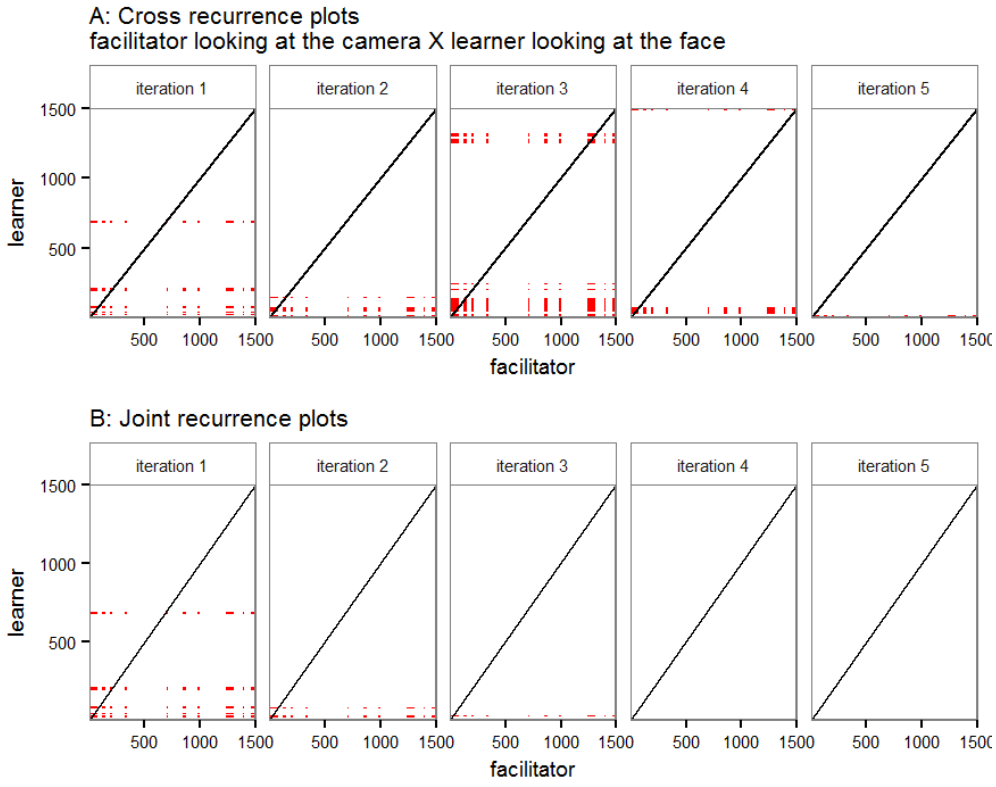


Figure 7-7. A: Cross recurrence plots (CRP) produced from the time series of the facilitator looking at the camera (horizontal axis) paired with the time series of learner_3 looking at the facilitator's face while watching the demonstration videos (vertical axis). The line of synchrony, in which the time stamp of the two time series coincide, is shown in black, and cross recurrences are shown in red. These are the same cases shown in **Figure 7-4**. The learner was in the condition FACE_AUDIO. B: Joint recurrence plots (JRP) produced from the CRPs are shown in A. For each iteration, the JRP is produced by multiplying the CRP from that iteration with all previous ones, thus leaving only those recurrences that occur in all of them. For iteration 1, the CRP is used. The number of recurrences drops dramatically and, from iteration 3, the JRPs are blank.

5) Diagonal-wise cross recurrence profiles

In the cross recurrence plots, the main diagonal or line of synchrony (slope = 1 and passing through the origin) compares the states exhibited by the facilitator and the learner at exactly the same time, while diagonals parallel to it (above or below the line of synchrony) compare the states of both participants at increasingly diverging time lags. Therefore, one way to visualize the degree of synchronization between two time series is by plotting the proportion of cross recurrence points that obtain at the line of synchrony and at each lag within a range of interest, here defined as from -3s to +3s. The shape of the resulting curve indicates the degree of synchrony between the facilitator and the learner. Specifically, a curve that peaks at around lag = 0 indicates that both time series tended to have the same labels at the same time, i.e. both participants tended to exhibit corresponding behavioural states at the same time. A curve that peaks at some other lag other than zero indicates that one time series tended to follow the other after some delay, with the delay indicated by that lag. A flat curve indicates that the sequence of states of the systems were not synchronized.

The resulting plots are called diagonal-wise cross recurrence profiles (Marwan et al., 2007). Two sets of such profiles were produced (**Figure 7-11**). One set examines synchronization between the facilitator manipulating the puzzle pieces and the learners looking at those pieces. The other set examines synchronization between the facilitator looking at the camera and learners looking at the facilitator's face on the screen. I report them to illustrate and provide qualitative evidence of synchronization. In addition, I report statistical models of RR and DET extracted from the respective cross recurrence plots, predicted by success, as described below.

Statistical analysis

To account for the hierarchical data structure that results from the mixed design, in the models reported below *condition* was modelled as a between-participant parameter with fixed effect, *iteration* was modelled as a within-participant parameter with fixed effect (repeated measures), and both the *indicator of puzzle* and the *indicator of participant* were modelled with varying intercepts (also called random effects). All analyses were carried out in `R` version 3.1.3 (R Core Team, 2016). Regression models were fit in STAN (B. Carpenter et al., 2017) using the convenient `map2stan` function from the `rethinking`

package (McElreath, 2016). STAN is a tool for implementing Bayesian multilevel models using Markov Chain Monte Carlo (MCMC) simulation, which approximates the posterior distribution for unknown parameters (Gelman et al., 2014; McElreath, 2016). Fitted Bayesian models provide samples from the posterior distribution of the parameters for which summary statistics are computed and reported (such as mean, credible intervals, differences, or the proportion of positive values). In all models, I used weakly informative priors to avoid unreasonable parameter values while still allowing the model to estimate a wide range of sensible values (Gelman et al., 2014), and I report posterior distributions estimated from at least 3,000 samples after warmup.

I first investigated if the situations created by the experimental manipulation might indeed be generally characterized as learning that is facilitated by the facilitator's behaviour (i.e. 'social learning'). To this end I fitted hierarchical logistic models with logit link to the performance data from participants in all six conditions. The outcome was the probability of success in each trial and the predictors were condition (coded as dummy variables), iteration, and the interaction between them. The model also included indicators of task and participant as varying intercepts (fully crossed random effects). No participant solved any of the tasks during the baseline test, and therefore there was no need to include the baseline score as a covariate, as was initially intended.

To examine whether the proportion of time learners spent looking at the puzzle pieces and at the facilitator's face varied across conditions, I fitted a hierarchical generalised linear model to the data. Because proportions are bounded between 0 and 1, I modelled these data using a Beta likelihood and logit link. The predictor was condition (coded as dummy variables) and I used indicators of label ('face' and 'pieces') such that the model would estimate the parameters for each separately. The model also included varying intercepts of task and participant. I also fitted a variation of this model adding iteration and its interaction with condition as predictors and compared them using the Widely Applicable Information Criterion or WAIC (Gelman et al., 2014; McElreath, 2016). There was no effect of iteration and WAIC indicated that adding iteration did not improve prediction accuracy. I therefore report the simpler model.

To further investigate if the way in which learners coordinated their eye movements with respect to the facilitator's manipulation varied across conditions, I fitted hierarchical

linear models to the data indicating coordination between the time series of *facilitator manipulating the puzzle pieces* paired with each time series of *learner looking at the pieces*. In each model, the outcome was one of the measures of coordination (RR, DET, or L, computed from either the cross recurrence plots or the joint recurrence plots) and the predictors were condition, iteration, and their interaction. The models also included varying intercepts for task and participant.

The observed values of RR extracted from the cross and joint recurrence plots computed from the times series of *facilitator manipulating the puzzle pieces* and *learner looking at the pieces* were close to zero (range: 1.79% - 10.90%), which reflected in some negative values being predicted by the fitted linear models. The posterior validation checks in these cases were good but since RR is a proportion and negative values do not make sense, I re-fitted the models which had RR (from CRP and from JRP) as the outcome using Beta distribution and logit link. These models converged, and mixing was good, but the posterior validation checks were less satisfactory than the linear models. Therefore, I opted to report the linear models, which make sense overall except for the caveat that RR is a proportion and therefore bounded within the unit interval.

The qualitative assessment of the diagonal-wise cross recurrence plots, which will be reported below, suggested a possible association between how learners looked at the facilitator's face and success at solving the task. I therefore also modelled the values of RR and DET extracted from the cross recurrence plots produced for the time series of *facilitator looking at the camera* paired with each time series of *learner looking at the facilitator's face* using hierarchical generalized linear models with Beta likelihood and logit link. In these models the outcome was either RR or DET and the only predictor was success in the trial. The models used indicators of condition to estimate effects for each separately. These models included varying intercepts for participants but not of trials since these could not be well estimated (as indicated by poor mixing of independent chains and low number of effective samples in the posterior distribution, in the models which included them).

Finally, to examine the association between coordination and learning, I used the framework of model comparison with WAIC and fitted hierarchical logistic models with logit link to the performance and coordination data. The guiding question in this analysis

is whether each of the six variables that quantify the coordination between the time series of *facilitator manipulating the puzzle pieces* and each time series of *learner looking at the pieces* were predictive of performance.

To conduct this analysis, I first modelled the probability of success predicted by condition, iteration, and the interaction condition X iteration. This 'base model' also included varying intercepts for task and participant but did not include any coordination variable. Then, for each of the six coordination variables, I produced two models with a structure that builds upon the base model. The first included the additive effect of the coordination variable and therefore asks if there is a statistic association between performance and coordination, after we take into account the effects of condition and iteration. The second model also included the interaction condition X coordination and therefore allow the effect of coordination (if there is any) to vary across conditions. Relatively lower values of WAIC are indicative of better predictive accuracy. Akaike weights are rescaled values of WAIC in which a total weight of 1 is partitioned among the models under consideration, thus indicating relative predictive accuracy among the models being compared (McElreath 2016). Comparing the WAIC and Akaike weight scores indicates whether adding each of the coordination variables (and the interaction with condition) improves prediction accuracy relative to the base model. The values of the coordination variables were standardised or z-scored (i.e. subtracted from the mean and divided by the standard deviation) to be included in the models.

Evaluating evidence of effects

All STAN models converged, and mixing was good. Unless otherwise indicated, I report the mean and 95% credible interval of the estimated parameters from the fitted models. In this section, strong evidence for an effect corresponds to a 95% central credible interval which excludes 0, and weak evidence corresponds to cases where the 95% central credible interval includes 0, but the 90% central credible interval does not. I sometimes report the probability that the model estimates a positive parameter value to provide further evidence of an effect. When evidence of an effect comes only from such a probability but not from the credible intervals (because even the 90% CI includes zero), I consider this to be very weak evidence. Unless noted, I report posterior estimates computed marginal

of task and participant, i.e. averaging across the variation introduced by their varying intercepts.

7.3 Results

7.3.1 Was learning influenced by the facilitator?

No participant solved any of the tasks during the baseline test. **Table 7-1**, **Table 7-2** and **Figure 7-8** report the model of the probability of success predicted by condition, iteration, and their interaction. In all conditions, the probability of success in the first trial after the baseline test was very low, suggesting that participants started from a similar level of skill and that the puzzles were relatively challenging. There was strong evidence that learning (improved performance with practice) was influenced by condition. This is indicated visually by how the curves showing simulated probabilities vary in steepness across conditions (**Figure 7-8**) and, formally, by the estimated effects of iteration (which indicate learning rate) and probabilities of success reported in **Table 7-1**, as well as by the differences in the effect of iteration across conditions reported in **Table 7-2**.

Table 7-1. Estimated mean and 95% credible intervals of the effect of iteration on the probability of success across conditions. Also reported are the estimated mean and 95% credible interval of the probability of success in the first and in the last trial.

Condition		Effect of iteration (mean and 95% CI)		probability of success at trial 1	probability of success at trial 5
		Estimate	Odds ratio		
noVIDEO noAUDIO		0.82 [0.08, 1.68]	2.27 [1.09, 5.39]	.03 [0, .30]	.14 [0, .89]
noVIDEO AUDIO		0.84 [0.30, 1.44]	2.32 [1.35, 4.23]	.05 [0, .53]	.28 [0, .97]
noFACE noAUDIO		0.91 [0.52, 1.35]	2.50 [1.69, 3.84]	.07 [0, .66]	.42 [.01, .99]
FACE noAUDIO		0.58 [0.19, 0.98]	1.79 [1.21, 2.67]	.12 [0, .85]	.36 [.01, 0.98]
noFACE AUDIO		1.81 [1.29, 2.38]	6.10 3.65, 10.76]	.06 [0, .68]	.88 [.30, 1]
FACE AUDIO		1.49 [1.01, 2.05]	4.44 [2.75, 7.74]	.06 [0, .59]	.74 [.08, 1]

To examine how learning varied across conditions in more detail, I first averaged the effects estimated for the relevant sets of conditions in each posterior simulation (thus producing a vector of average estimates) and then computed the difference between the sets (**Table 7-2**). Compared to engaging with the puzzles individually (with no influence from the facilitator), listening to the facilitator's speech alone without watching the corresponding videos did not lead to a faster learning rate (comparison 1) and neither did watching the demonstration videos without listening to the corresponding audio (comparison 2). In contrast, participants who watched the videos *and* listened to the facilitator's speech learned faster compared to those who did not watch the videos (comparisons 3 and 4), and also compared to those who watched the videos without the audio (comparisons 5 and 6). There was no evidence of a difference in learning rate between those watching the video with the facilitator's face visible, compared to blurred (comparisons 7 and 8). In fact, there is a weak trend in the opposite direction. The probability that the model estimates a greater effect of iteration in condition noFACE_AUDIO than FACE_AUDIO is .81, and that it estimates a greater effect in condition noFACE_noAUDIO than FACE_noAUDIO is .88. Thus, there is some suggestion that allowing participants to see the facilitator's face in the demonstration videos proved detrimental to learning in this case.

Table 7-2. Estimated difference in the effect of iteration on the probability of success across sets of conditions. Unless indicated, reported values are mean and 95% credible intervals.

comparison	Conditions being compared		Difference in the effect of iteration (set 2 – set 1)
	set 1	set 2	
1	noVIDEO_noAUDIO	noVIDEO_AUDIO	0.02 [-0.97, 0.93]
2	noVIDEO_noAUDIO	noFACE_noAUDIO, FACE_noAUDIO	-0.07 [-0.91, 0.79]
3	noVIDEO_noAUDIO	noFACE_AUDIO, FACE_AUDIO	0.83, 90% CI [0.10, 1.57]
4	noVIDEO_AUDIO	noFACE_AUDIO, FACE_AUDIO	0.81 [0.16, 1.49]
5	noFACE_noAUDIO	noFACE_AUDIO	0.89 [0.25, 1.55]
6	FACE_noAUDIO	FACE_AUDIO	0.91 [0.28, 1.54]
7	noFACE_noAUDIO	FACE_noAUDIO	-0.34 [-0.91, 0.21]
8	noFACE_AUDIO	FACE_AUDIO	-0.32 [-1.07, 0.34]

In sum, neither listening to the facilitator's speech alone (without watching the demonstration) or watching the demonstration alone (without listening to the facilitator's speech) improved learning rate, compared to engaging in the task individually. In contrast, listening to the facilitator's speech while watching the corresponding demonstration greatly facilitated learning. In addition, there was some evidence suggesting that seeing the facilitator's face, compared to face blurred, was relatively detrimental to learning.

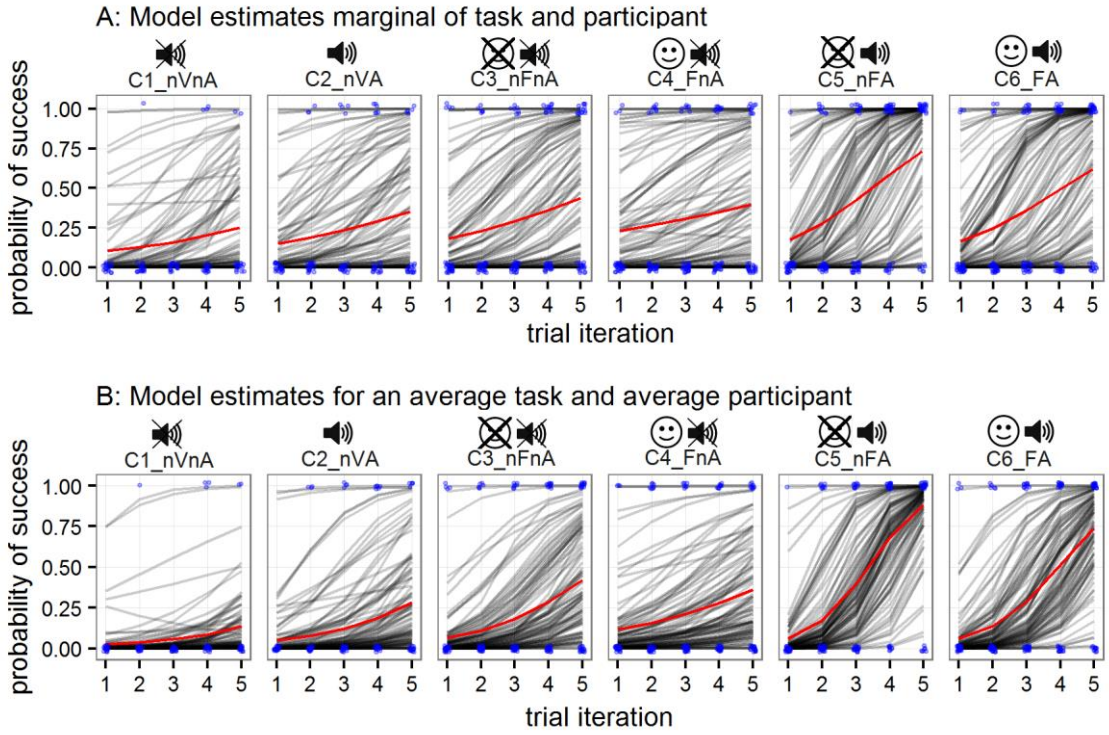


Figure 7-8. Plots show 100 simulations (black) and the mean predictions (red) from the fitted model of performance as a function of condition, iteration, and their interaction across all conditions. A: simulations marginal of task and participant, i.e. considering the variation introduced by their varying intercepts. B: simulations for an average task and average participant, i.e. without considering the variation introduced by the varying intercepts of task and participant. The observed performance data are shown in blue (0: no success, 1: success), with their position jittered to avoid overlap. C1_nVnA: no video and no audio. C2_nVA: no video, with audio. C3_nFnA: video with noFACE_noAUDIO. C4_FnA: video with FACE_noAUDIO. C5_nFA: video with noFACE_AUDIO. C6_FA: video with FACE_AUDIO.

7.3.2 Did condition affect relative allocation of visual attention?

Figure 7-9 reports the model of the proportion of fixation time participants directed at the facilitator's face and at the puzzle pieces, predicted from condition. Overall, participants looked for much longer at the puzzle pieces compared to the facilitator's face (difference in the mean estimates averaged across conditions: 74.1% [40.3%, 90.1%]).

There was no effect of condition on the time spent looking at the pieces. In contrast, there was strong evidence that participants looked more at the facilitator's face when it was visible compared to blurred (difference in the mean estimates between FACE_noAUDIO and noFACE_noAUDIO: 3.14% [0.5%, 10.3%]), between FACE_AUDIO and noFACE_AUDIO: 5.6% [0.8%, 17.9%]). There was no evidence of a difference in the mean estimates between the two conditions in which the face was blurred. Comparing the two conditions in which the face was visible, there was weak evidence that participants spent more time looking at his face when the audio was played (difference between FACE_AUDIO and FACE_noAUDIO: 2.9%, 90% CI [0.2%, 8.0%]]). In sum, learners looked more at the facilitator's face when it was visible compared to blurred, and even more when they could also listen to what he was saying.

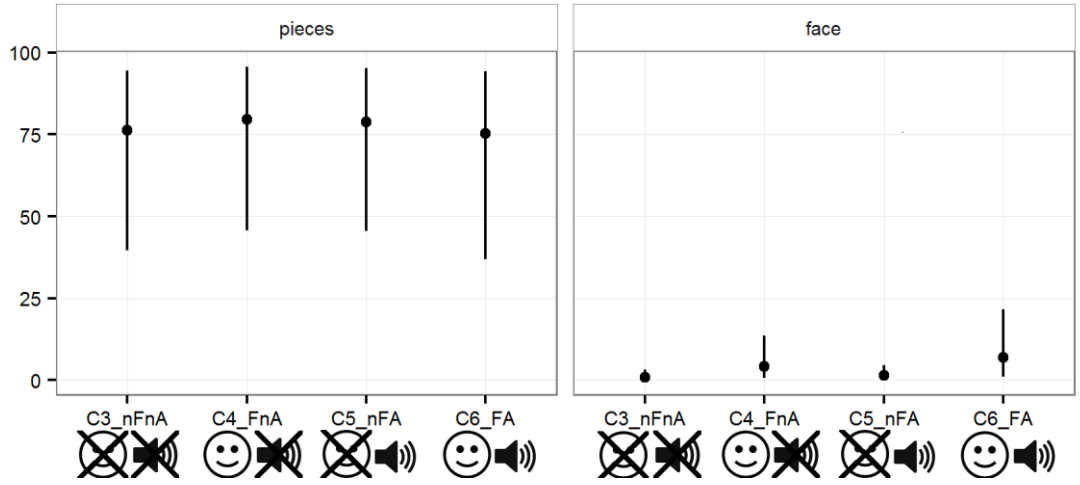


Figure 7-9. Estimates and 95% credible interval of the proportion of fixation time participants spent looking at the facilitator's face and at the puzzle pieces, across the four experimental conditions under scrutiny. C3_nFnA: noFACE_noAUDIO. C4_FnA: FACE_noAUDIO. C5_nFA: noFACE_AUDIO. C6_FA: FACE_AUDIO.

7.3.3 Did condition affect the coordination between the learners' visual attention and the facilitator's behaviour?

I first focus on the analyses of the time series representing how the facilitator manipulated the puzzle pieces paired with each time series representing how learners looked at the pieces. **Table 7-3** and **Figure 7-10** report the models of the six different coordination variables, in which the predictors were condition, iteration, and their interaction. The estimates of the coordination variables overlap considerably across conditions and across iterations, and there was no evidence of an effect of condition on the mean estimates,

averaged across iterations, of any of the coordination variables. With regard to the variables computed from the cross recurrence plots, there was strong evidence that DET and L increased across iterations in the condition noFACE_AUDIO.

With respect to the joint recurrence plots (JRPs), there was no evidence of an effect of condition on the mean estimates of RR and DET, averaged across iterations. There was evidence of an effect of condition on values of L. Comparing the two conditions with the facilitator's face blurred, there was weak evidence that the mean estimates of L (from JRPs), averaged across iterations, were lower when the audio was played (difference between the estimates of conditions noFACE_AUDIO and noFACE_noAUDIO: -5, 90% CI [-11, -0.1]). This suggests that listening to the audio might have led learners to look at different pieces in different trials. In addition, comparing the two conditions in which the audio was played, there was strong evidence that the mean estimates of L were higher when the facilitator's face was visible compared to blurred (difference between the mean estimates of conditions FACE_AUDIO and noFACE_AUDIO: 8 [1, 13]). This suggests that seeing the facilitator's face, compared to face blurred, motivated learners to stay synchronized with the facilitator's actions more similarly across trials. Each cross recurrence point corresponds to approximately 25ms, thus, for example, the estimated difference of 5 corresponds to about 125ms.

Table 7-3. Estimated values for the effect of iteration on the different coordination variables: RR, DET, and L, computed from the cross recurrence plots (CRP) and the joint recurrence plots (JRP). Reported values are means and 95% credible intervals. Values indicating strong evidence of an effect are in bold and shaded to facilitate reading.

RP	Variable	Conditions			
		noFACE noAUDIO	**FACE noAUDIO**	**noFACE AUDIO**	**FACE AUDIO**
CRP	**RR**	0.04 [-0.04, 0.11]	0.02 [-0.06, 0.11]	0 [-0.08, 0.09]	-0.01 [-0.09, 0.07]
	DET	0.36 [-0.56, 1.28]	0.79 [-0.24, 1.75]	**1.37 [0.25, 2.51]**	0.64 [-0.28, 1.58]
	L	0.48 [-0.65, 1.56]	0.4 [-0.79, 1.62]	**1.58 [0.34, 2.86]**	0.53 [-0.57, 1.62]
JRP	**RR**	**-0.60 [-0.70, -0.50]**	**-0.64 [-0.74, -0.53]**	**-0.60 [-0.72, 0.49]**	**-0.68 [-0.79, -0.58]**
	DET	**-6.71 [-8.00, -5.46]**	**-6.66 [-8.03, -5.34]**	**-6.5 [-7.98, -4.99]**	**-5.73 [-6.96, -4.49]**
	L	**-4.16 [-5.1 -3.18**	**-4.52 [-5.54, -3.44]**	**-3.45 [-4.57, -2.27]**	**-3.74 [-4.71, -2.75]**

Because of the way they are computed, described above, JRPs produced by multiplying more (compared to fewer) CRPs will tend to have lower values of RR. JRPs for later (compared to earlier) trials are produced by multiplying more CRPs. For example, the JRP for trial 2 is produced by multiplying the CRPs for trials 1 and 2, whereas the JRP for trial 5 is produced by multiplying the CRPs for trials 1, 2, 3, 4, and 5. It is therefore expected that JRPs from later trials will have lower values of RR (see **Figure 7-5** and **Figure 7-7** for examples). Consistent with this, there was strong evidence that RR, DET, and L were negatively associated with iteration in all four conditions, but there was no evidence of a difference in the magnitude of this effect across conditions in any of these cases.

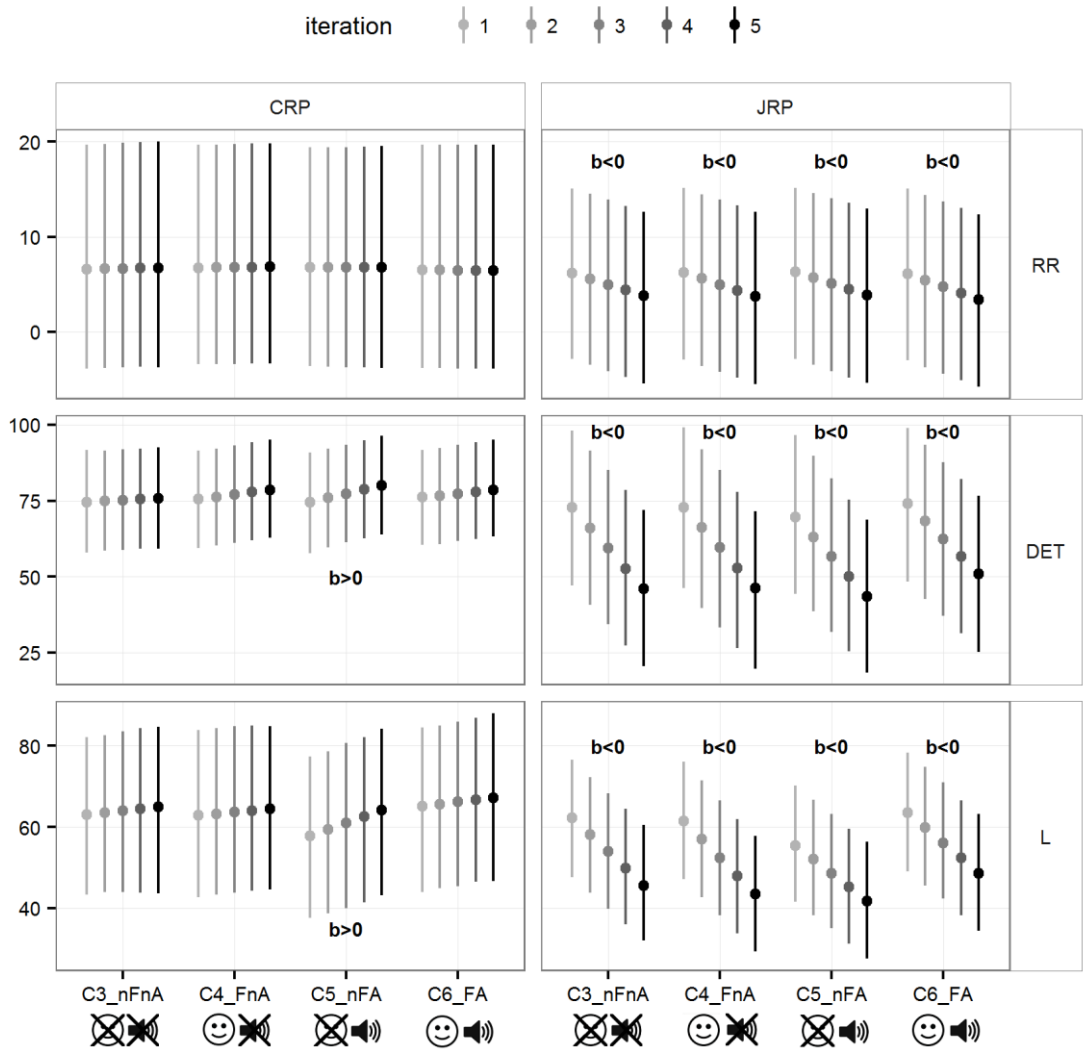


Figure 7-10. Estimated mean and 95% credible intervals of RR, DET, and L, for iterations 1 (light grey) to 5 (black), marginal of task and participant, across the experimental conditions under scrutiny. The labels b<0 (or b>0) indicate cases where there is strong evidence of a negative (or positive) effect of iteration on the coordination variable. C3_nFnA: noFACE_noAUDIO. C4_FnA: FACE_noAUDIO. C5_nFA: noFACE_AUDIO. C6_FA: FACE_AUDIO.

Finally, **Figure 7-11** shows the two sets of diagonal-wise cross recurrence profiles computed to illustrate and examine synchronization between learner and facilitator qualitatively. I start with set A, which examines synchronization between the time series of *facilitator manipulating the puzzle pieces* paired with each time series of *learner looking at the pieces*. The peak around lag = 0 means that most recurrence points in the resulting cross recurrence plots fall on or close to the line of synchrony, as evident in the examples in **Figure 7-5**. In other words, in all trials, participants showed a strong

tendency to be looking at precisely those puzzle pieces being manipulated by the facilitator at any moment. The apparent similarity across conditions, and also between successful and unsuccessful trials, suggests there were no qualitative differences in how synchronized participants' eye movements were relative to the pieces being manipulated by the facilitator. This is consistent with the lack of evidence of an effect of condition on the coordination variables extracted from cross recurrence plots.

I now focus on the analyses of the time series representing how the facilitator looked at the camera paired with each time series representing how learners looked at the facilitator's face. Set B in **Figure 7-11** examines synchronization in this case. The overall low values of RR (compared to set A) indicate that the combination of the states examined here was substantially less recurrent over time in each trial than the states examined in set A. There are only very few curves with values above zero in conditions noFACE_noAUDIO and noFACE_AUDIO, indicating that learners never (or only rarely) looked at the facilitator's face when it was blurred. This is consistent with the evidence from the proportion of fixation time and fixation count, reported above. The fact that these curves do not show a clear peak indicates that, in those cases when learners did look at the facilitator's blurred face, they did so irrespective of the facilitator looking at the camera. This is evidence that the amount of blurring applied to the videos did indeed disrupt the optical structure which specified the facilitator looking at the camera, as intended.

In contrast, there are many more curves with values of cross recurrence above zero in condition FACE_noAUDIO, and even more in FACE_AUDIO. This indicates that some learners did, at least in some trials, look at the facilitator's face when it was clearly visible, and they did so more when the audio was played. The fact that the curves in the conditions with face visible show a peak close to lag = 0 indicate that, in those cases when learners did look at the facilitator's face, they did so at about those moments when the facilitator looked at the camera. Moreover, the fact that these curves tend to peak at slightly positive lags indicates that the facilitator was leading this behaviour, i.e. the state *facilitator looking at the camera* tended to occur first, closely followed by the state *learner looking at the facilitator's face*. This suggests learners were responding to the

facilitator looking ‘at them’ or, more precisely, to the sudden appearance, in their field of view, of optical structure specifying a pair of eyes looking in their direction.

Visual inspection of the profiles suggests that, on average, the values of RR were higher, and the profile peaks were also higher, in successful compared to unsuccessful trials. This is particularly suggestive in condition FACE_AUDIO. This qualitative suggestion was confirmed formally by the model of RR predicted by success, which provided strong evidence of a positive association in condition FACE_AUDIO (b = 0.26 [0.01, 0.54]). There was also a weak trend of a positive association in condition FACE_noAUDIO, suggested by the probability that the model estimates a positive effect (p = .77). The models of DET predicted by success provided only weak evidence of a trend in both conditions in which the face was visible, supported by the probability that the model estimates a positive effect (FACE_noAUDIO: p = .86; FACE_AUDIO: p = .83).

When reporting the relative effect of iteration on performance across conditions above, I provided some (relatively weak) evidence suggesting that seeing the facilitator’s face in the demonstration videos was detrimental to learning. Here I provide evidence that looking back at the facilitator’s face was associated with success. This might sound contradictory and therefore it is sensible to clarify. Consider that the effects of iteration on success are smaller in condition FACE_noAUDIO compared to noFACE_noAUDIO, and also smaller in condition FACE_AUDIO compared to noFACE_AUDIO (respectively, comparisons 7 and 8 in **Table 7-2**). This difference is also illustrated by the learning curves depicted in **Figure 7-8**. This comparison is among participants that find themselves *in different learning conditions*. Here the comparison is among participants that find themselves *within the same learning conditions* (either FACE_noAUDIO or FACE_AUDIO).

Thus, on the one hand, I provided evidence that the presence of the optic structure specifying the facilitator’s face (compared to a pixelized blob) in the visual array was somewhat detrimental to learning. This is possibly because the face momentarily distracts the learners’ attention from the actions being performed on the pieces. On the other hand, given that the face is visible, successful participants tended to be those responsive to the facilitator’s glancing at them (more precisely, glancing at the camera during video

recording). Since the structure of the facilitator's face is unrelated to the solution of the tasks, I suggest that the behaviour of 'being looked at' by the facilitator influences the motivational state of the learners, and that this effect is even stronger when they can also hear the facilitator speaking 'to them'. Assuming that the tasks were not so difficult such that all participants could eventually learn how to solve them, this motivational influence might give these learners an advantage. This effect is possibly triggered by the participants' prior experience in similar situations. On the other hand, it is possible that these participants were more motivated regardless and that their behaviour of 'looking back' is capturing their intrinsic motivation.

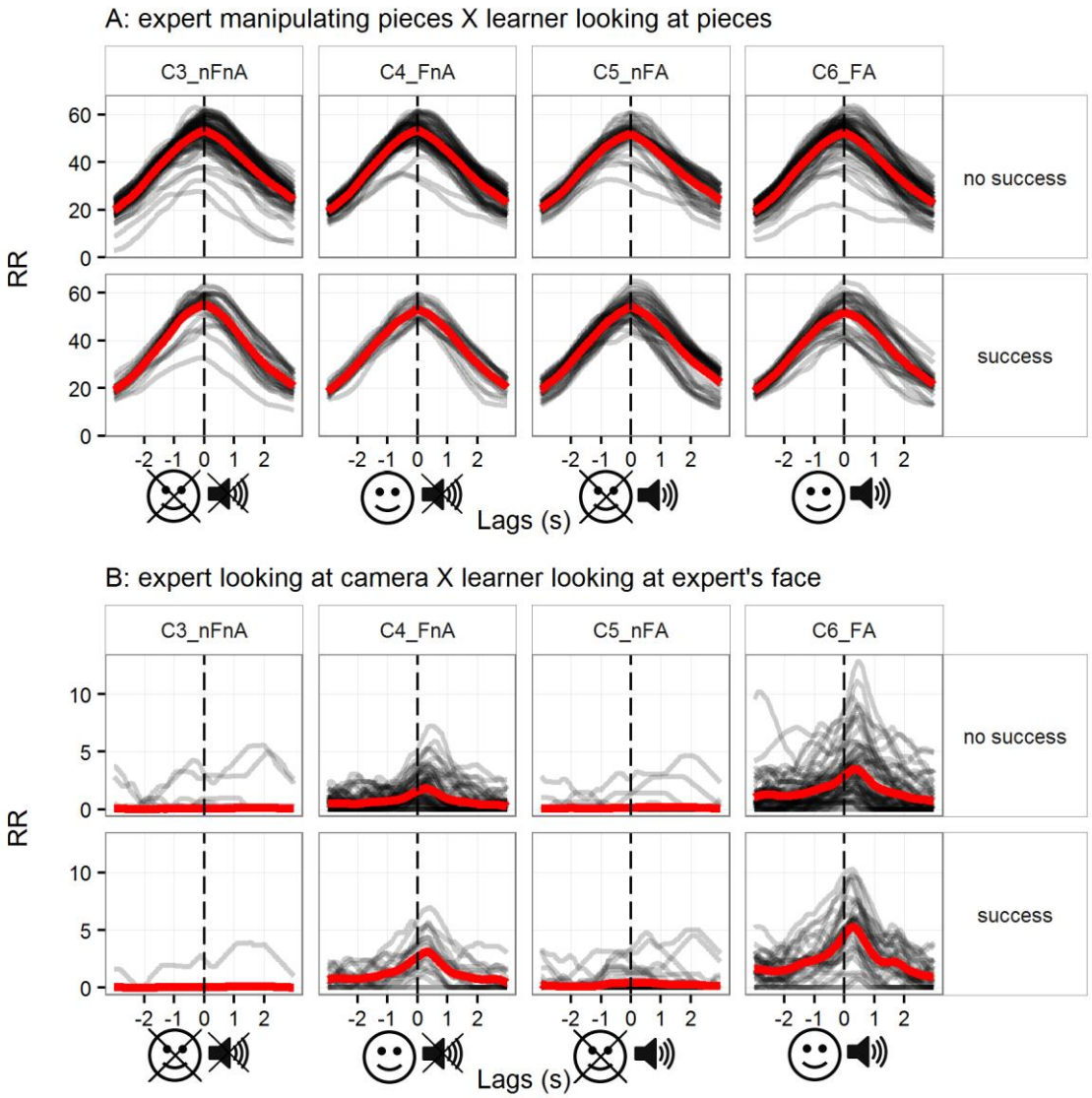


Figure 7-11. Diagonal-wise cross recurrence profiles computed for the observed data (black), where each curve represents one trial, and the mean across trials (red). The values of each curve indicate the proportion of recurrence points (cross recurrence rate) that fall within a range of lags (from -3s to 3s) around the main diagonal or line of synchrony (lag = 0) in the corresponding cross recurrence plot. A: This set examines synchronization between the time series of the facilitator manipulating the puzzle pieces and the time series of the learners looking at those pieces as they watch the demonstration videos. B: This set examines synchronization between the time series of the facilitator looking at the camera and the time series of the learners looking at the facilitator's face. Conditions are C3_nFnA: noFACE_noAUDIO; C4_FnA: FACE_noAUDIO; C5_nFA: noFACE_AUDIO; C6_FA: FACE_AUDIO.

7.3.4 Was learning associated with measures of coordination?

Table 7-4 reports the model comparison conducted for the models of performance to begin examining the association between coordination and learning. According to the WAIC and Akaike weight values, including the coordination variables extracted from cross recurrence plots did not improve the prediction accuracy relative to the base model with predictors condition, iteration, and their interaction, and these models will therefore not be further analysed. On the other hand, including coordination variables extracted from joint recurrence plots did improve prediction accuracy relative to the base model. More specifically, including RR improved model accuracy but including the interaction between condition and RR did not improve it any further. With respect to DET and L, including both the coordination variable and its interaction with condition improved model accuracy. I therefore focus on these three models below.

Table 7-5 reports the estimated effects of iteration on the probability of success in these models. As expected, these effects are similar to those reported in **Table 7-1**. **Table 7-6** reports the estimated effects of coordination on the probability of success. To aid interpretation of what the estimates imply, I present a set of three figures corresponding to each model: The model with RR as a covariate is shown in **Figure 7-12**, the model with DET in **Figure 7-13**, and the model with L in **Figure 7-14**. These figures show model predictions of the probability of success as a function of condition and iteration, computed for three different values of the respective coordination variable: -2, 0, and 2. Recall that these variables were standardised to be included as covariates. Therefore, 0 corresponds to the observed mean value, -2 corresponds to the value 2 standard deviations below the mean (i.e. a low value), and 2 corresponds to the value 2 standard deviations above the mean (i.e. a high value). Note that these figures are similar to the model of performance shown above in **Figure 7-8**, which included only condition, iteration, and their interaction, as predictors. At the same time, they present their individual assumptions, predictions, and uncertainties.

Table 7-4. Model comparison for the models of performance (including, or not including, coordination variables as predictors) using information criteria. WAIC: Widely Applicable Information Criterion; dWAIC: difference between each WAIC and the lowest WAIC; weight: Akaike weight (rescaled WAIC); SE: standard error of WAIC estimates; dSE: standard error of the difference in WAIC between each model and the top-ranked model. The final models, chosen for interpretation and reporting, are shown in grey.

Recurrence plot	Model	WAIC	dWAIC	weight	SE	dSE
CRP	Base model	323.9		0.42	25.27	
	+ RR	323.9	0	0.41	25.27	1.17
	+ RR + condition X RR	325.7	1.9	0.17	26.06	5.28
	Base model	323.9		0.54	25.27	
	+ DET	324.3	0.5	0.43	25.34	0.63
	+ DET + condition X DET	329.6	5.8	0.03	26.23	4.4
	Base model	323.9		0.58	25.27	
	+ L	324.6	0.7	0.4	25.37	0.67
	+ L + condition X L	331.2	7.4	0.01	26.29	4.33
JRP	+ RR	319.5		0.6	25.32	
	+ RR + condition X RR	320.7	1.2	0.33	26.29	5.49
	Base model	323.9	4.4	0.07	25.27	4.27
	+ DET + condition X DET	307.3		1	25.16	
	Base model	323.9	16.6	0	25.27	9.24
	+ DET	325.5	18.2	0	25.55	9.33
	+ L + condition X L	309.4		1	26.16	
	Base model	323.9	14.4	0	25.27	9.22
	+ L	324.6	15.2	0	25.2	9.09

Table 7-5. Estimated values for the effect of iteration on the probability of success from the fitted models of success predicted by condition, iteration, their interaction, and one of the coordination variables extracted from joint recurrence plots. Values indicating strong evidence of an effect are in bold and shaded in dark grey to facilitate reading.

Coordination variable (from JRPs)	Condition		Effect of iteration on the probability of success (mean and 95% CI)	
			Estimate	Odds ratio
RR	noFACE noAUDIO		**1.11 [0.55, 1.63]**	**3.17 [1.72, 5.08]**
	FACE noAUDIO		**0.81 [0.26, 1.37]**	**2.34 [1.18, 3.74]**
	noFACE AUDIO		**2.00 [1.34, 2.67]**	**7.83 [3.37, 13.62]**
	FACE AUDIO		**1.95 [1.26, 2.65]**	**7.56 [3.09, 13.47]**
DET	noFACE noAUDIO		**1.39 [0.70, 2.12]**	**4.33 [1.74, 7.79]**
	FACE noAUDIO		0.15 [-0.66, 0.91]	1.27 [0.43, 2.30]
	noFACE AUDIO		**2.70 [1.76, 3.68]**	**17.04 [4.71, 36.47]**
	FACE AUDIO		**1.44 [0.78, 2.15]**	**4.51 [1.82, 7.95]**
L	noFACE noAUDIO		**1.73 [0.98, 2.55]**	**6.18 [2.30, 11.84]**
	FACE noAUDIO		0.01 [-0.77, 0.77]	1.10 [0.39, 2.02]
	noFACE AUDIO		**2.20 [1.37, 3.07]**	**10.05 [3.15, 19.90]**
	FACE AUDIO		**1.58 [0.97, 2.28]**	**5.17 [2.20, 8.96]**

Table 7-6. Estimated values for the effect of coordination variables computed from joint recurrence plots (JRP) on the probability of success across conditions, computed for the models chosen following the model comparison. The probability that the models estimate a positive effect given the data is also reported. Values indicating strong (or weak) evidence of an effect are in bold and shaded in dark (or light) grey to facilitate reading.

Coordination variable (from JRPs)	Condition	Effect of coordination on the probability of success (mean and 95% CI)	
		Estimate	Odds ratio
RR	**noFACE_noAUDIO, FACE_noAUDIO, noFACE_AUDIO, and FACE_AUDIO**	**0.91 [0.02, 1.78]**	**2.48 [1.02, 5.93]**
DET	**noFACE noAUDIO**	**1.14 [0.01, 2.29]**	**3.13 [1.01, 9.91]**
	FACE noAUDIO	-1.32 [-3.03, 0.45]	0.27 [0.05, 1.57]
	noFACE AUDIO	**2.14 [0.81, 3.68]**	**8.50 [2.25, 39.49]**
	FACE AUDIO	**-1.11 [-2.12, -0.17]**	**0.33 [0.12, 0.85]**
L	**noFACE noAUDIO**	**2.05 [0.82, 3.28]**	**7.76 [2.28, 26.45]**
	FACE noAUDIO	**-1.82 90% CI [-3.42, -0.24]**	**0.16 90% CI [0.03, 0.79]**
	noFACE AUDIO	**1.39 90% CI [0.02, 2.71]**	**4.00 90% CI [1.02, 14.96]**
	FACE AUDIO	-0.41 [-1.11, 0.30]	0.66 [0.33, 1.35]

There was strong evidence that RR was positively associated with probability of success, although the effect size is relatively small. This is indicated by the model predicting slightly higher probabilities of success for high (2) compared to low (-2) values of RR. Because this model does not include the interaction between condition and RR, the additive effect of RR is the same across all conditions. This model tells a very coherent story, for three reasons. First, the comparison between the conditions [noFACE_noAUDIO and FACE_noAUDIO] with conditions [noFACE_AUDIO and FACE_AUDIO] shows that learning was facilitated by listening to the audio instructions compared to not. Second, the comparison, within each condition, among different values of RR shows that learning was also facilitated by coordinating eye movements with

respect to the manipulations being demonstrated. Finally, comparing between conditions noFACE_noAUDIO and FACE_noAUDIO, as well as between noFACE_AUDIO and FACE_AUDIO, show that learning was slower when the facilitator's face was clearly visible compared to blurred.

The variables DET and L show a more complicated pattern. In the conditions where the facilitator's face was blurred, both DET and L were positively associated with probability of success. This suggests that, in these cases, learners who synchronized their eye movements more and for longer with the facilitator's actions learned faster. Regarding DET, the evidence of this effect is strong in conditions noFACE_noAUDIO and noFACE_AUDIO. Regarding L, the evidence of this effect is strong in condition noFACE_noAUDIO and weak in condition noFACE_AUDIO. So far this is consistent with the model of RR and with the general prediction that coordinating visual attention to the actions being demonstrated should be beneficial for learning. However, in the conditions where the facilitator's face was visible, DET and L were negatively associated with probability of success, as shown by the negative estimates reported in **Table 7-6**. With respect to DET, the evidence of this effect is strong in condition FACE_AUDIO but there is also a trend in condition FACE_noAUDIO, since here the model estimated a negative effect with probability p = .93, but the 90% credible interval included zero. Regarding L, the evidence of this effect is weak in condition FACE_noAUDIO but there is also a trend in condition FACE_AUDIO, since here the probability that the model estimated a negative effect was p = .88. I find this pattern surprising, and it is unclear how best to explain this result.

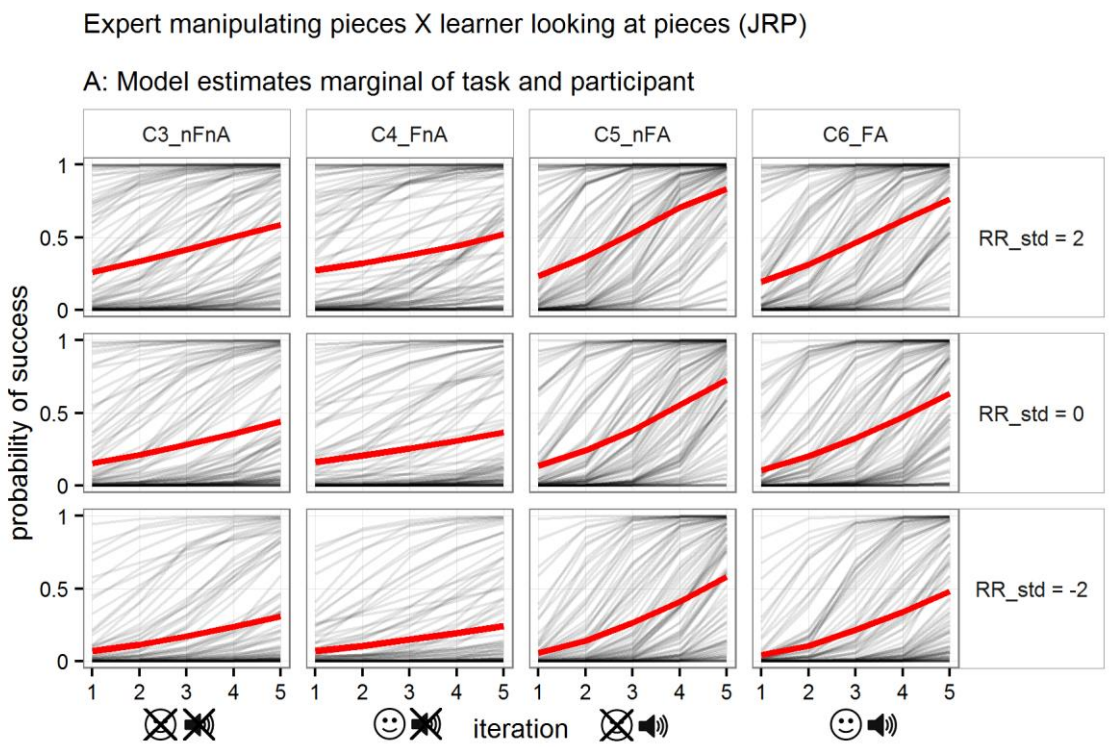


Figure 7-12. Estimated probability of success across conditions C3-C6 and across iterations, computed for low (-2), mean (0) and high (2) values of RR computed from joint recurrence plots. To see the effect of the coordination variable in each condition, the reader should compare the three plots in the same column. To see the effect of a given value of the coordination variable across conditions, the reader should compare the plots in the same row. Here the joint recurrence plots were computed from the time series indicating the pieces being manipulated by the facilitator and the time series indicating the pieces being foveated by the learner. To indicate uncertainty, reported values are 100 simulations from the fitted models. Conditions are C3_nFnA: noFACE_noAUDIO; C4_FnA: FACE_noAUDIO; C5_nFA: noFACE_AUDIO; C6_FA: FACE_AUDIO.

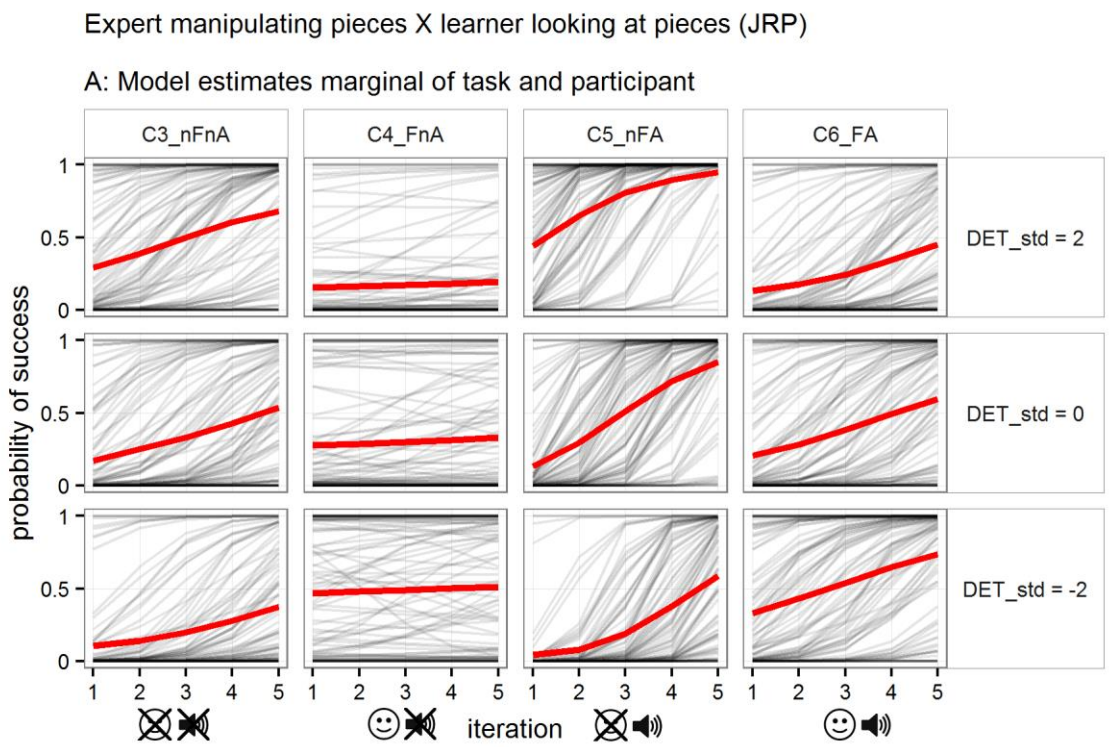


Figure 7-13. Estimated probability of success across conditions C3-C6 and across iterations, computed for low (-2), mean (0) and high (2) values of DET. computed from joint recurrence plots. To see the effect of the coordination variable in each condition, the reader should compare the three plots in the same column. To see the effect of a given value of the coordination variable across conditions, the reader should compare the plots in the same row. Here the joint recurrence plots were computed from the time series indicating the pieces being manipulated by the facilitator and the time series indicating the pieces being foveated by the learner. To indicate uncertainty, reported values are 100 simulations from the fitted models. Conditions are C3_nFnA: noFACE_noAUDIO; C4_FnA: FACE_noAUDIO; C5_nFA: noFACE_AUDIO; C6_FA: FACE_AUDIO.

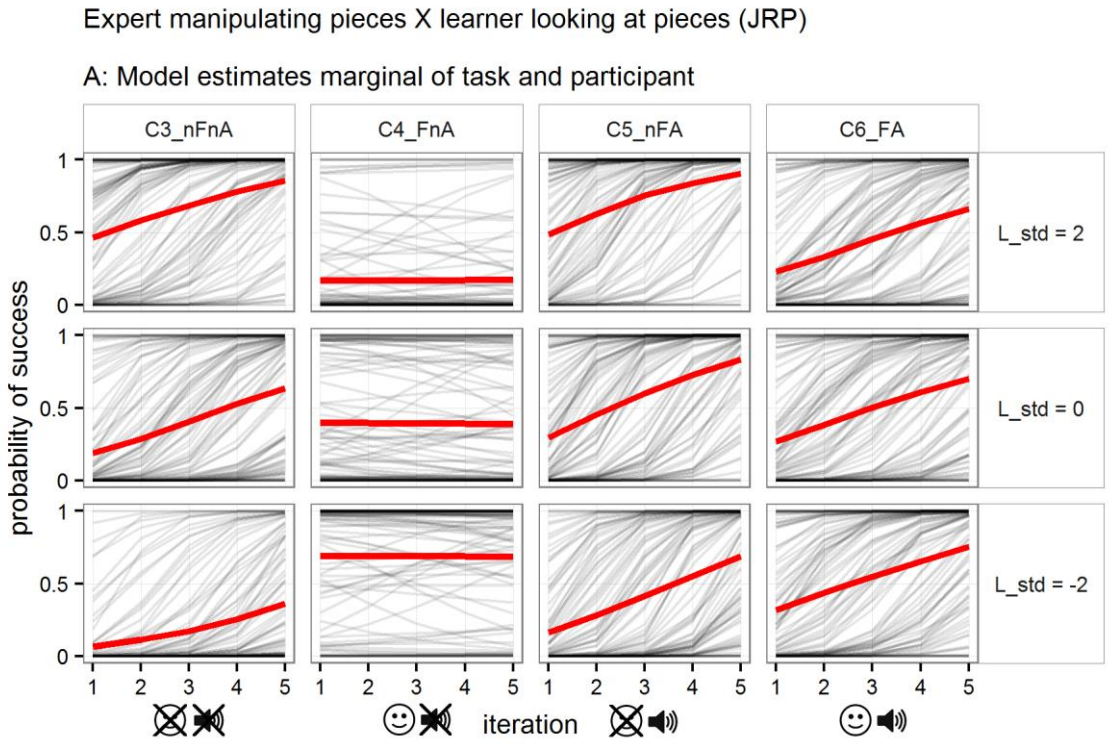


Figure 7-14. Estimated probability of success across conditions C3-C6 and across iterations, computed for low (-2), mean (0) and high (2) values of L. computed from joint recurrence plots. To see the effect of the coordination variable in each condition, the reader should compare the three plots in the same column. To see the effect of a given value of the coordination variable across conditions, the reader should compare the plots in the same row. Here the joint recurrence plots were computed from the time series indicating the pieces being manipulated by the facilitator and the time series indicating the pieces being foveated by the learner. To indicate uncertainty, reported values are 100 simulations from the fitted models. Conditions are C3_nFnA: noFACE_noAUDIO; C4_FnA: FACE_noAUDIO; C5_nFA: noFACE_AUDIO; C6_FA: FACE_AUDIO.

In an attempt to understand how this pattern arises from the data, I ran a post-hoc model with the same structure as above. The outcome was success and the predictors were condition, iteration, the interaction condition x iteration, coordination, and the interaction coordination x condition. As always, it also included varying intercepts for task and participant. However, this time I included the coordination variables RR, DET, and L, extracted from the CRPs produced by pairing the time series of the facilitator looking at the camera paired with each time series of the learners looking at his face. Thus, these models examine the association between 'mutual gaze' (or its equivalent in this situation) and learning. I used WAIC to compare each of the three models with the base model (which does not include the coordination variable). In the next paragraph, I use the subscripts (e.g. RR_{face}) for clarity.

According to this comparison, adding the coordination variable did not improve prediction accuracy and, examining the estimates, there was no evidence of an effect of RR_{face} on the probability of success in any condition. There was, however, some weak evidence of effects of DET_{face}, and L_{face} on success which varied across conditions, suggested by the probability that the model estimates a positive or negative effect. The posterior distribution of these effects is shown in **Figure 7-15**. The trend was the following. The effect of DET_{face} on success was *positive* in the two conditions in which learners could see the facilitator' face (FACE_noAUDIO and FACE_AUDIO), which are precisely the conditions for which the model above estimated a *negative* effect of DET_{pieces} on success. Moreover, the magnitude of the effect was higher in FACE_AUDIO than in FACE_noAUDIO with $p = .90$. Similarly, the effect of L_{face} was also *positive* in one of the conditions in which learners could see the facilitator's face (FACE_noAUDIO), for which the model above estimated a *negative* effect of L_{pieces}. This suggests that the presence of the facilitator's face affected how learners engaged in the task, even if the face in itself cannot inform the learners about the solution to the task.

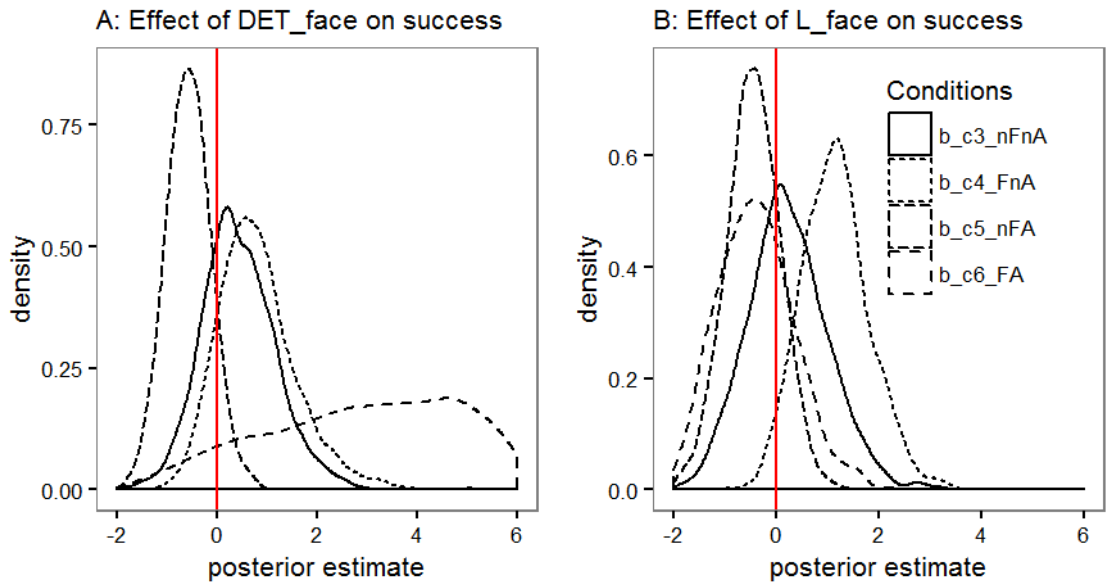


Figure 7-15. Posterior distribution of the estimated effect of DET (A) or L (B) on success, in each condition. In this model, DET and L were extracted from the CRPs produced by pairing the time series of the facilitator looking at the camera paired with each time series of the learners looking at his face. The vertical line indicates zero. The model tended to estimate a negative effect for condition noFACE_AUDIO, and a positive effect for the other conditions. The range of estimated values indicates the uncertainty of the model given the data and model assumptions. Conditions are C3_nFnA: noFACE_noAUDIO; C4_FnA: FACE_noAUDIO; C5_nFA: noFACE_AUDIO; C6_FA: FACE_AUDIO.

7.4 Discussion

Solving puzzles and watching videos as relational processes

Participants were invited to solve a set of construction puzzles. After a baseline test, participants watched videos in which the facilitator demonstrated the solution to the target task, and subsequently attempted to solve it themselves again. This procedure – watching a video and having a go at the task – was repeated five times for each of three different puzzles. Note that the potential learners did not interact with the facilitator directly. Participants received different video stimuli depending on condition. The facilitator's face was either visible or not, and the corresponding audio tracks were either played or not. I examined how the different stimuli affected learning and variables which capture the real-time behavioural coordination between potential learners and the information (*sensu* meaningful patterns of light and sound) made available by the videos. I also examined the association between coordination variables and learning. Following up on the previous chapter, this study was used as an opportunity to further explore what it means to view social learning from the perspective of developmental systems thinking and radical embodied cognitive science. **Figure 7-16** represents the learning situation produced by the study design, conceived from this perspective, schematically.

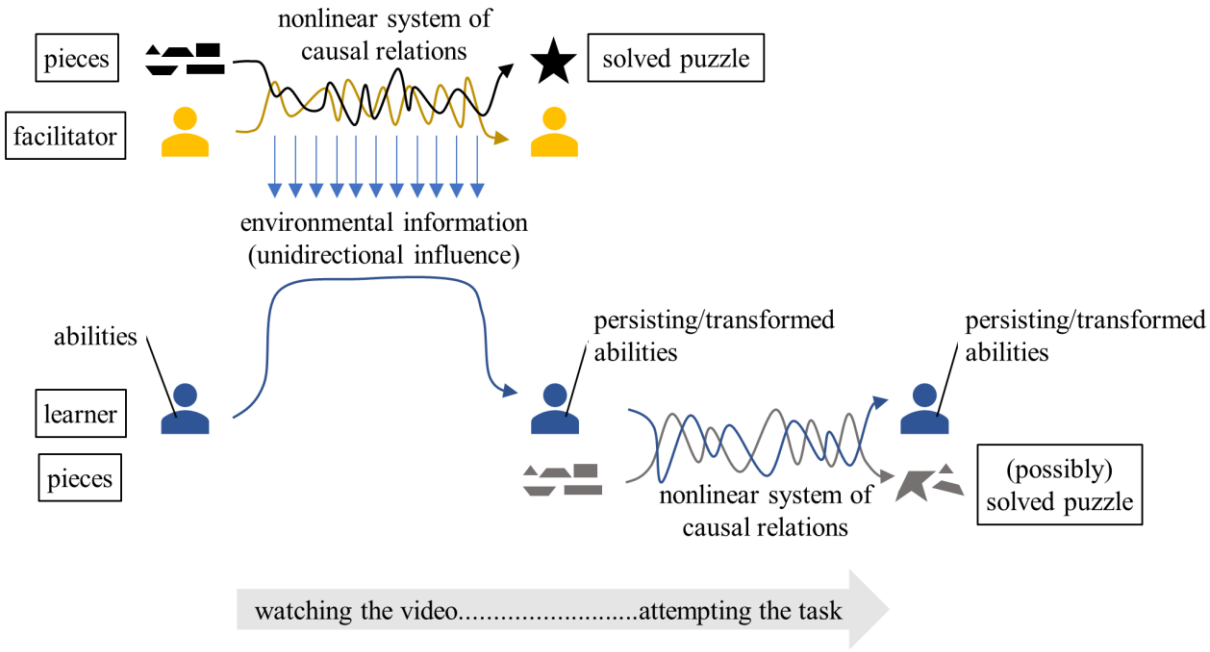


Figure 7-16. The learning situation conceived of as a relational developmental process. In each iteration, learners have the opportunity to sample and make sense of the environmental information produced by the

video showing the facilitator demonstrating how to solve the task. This is followed by an attempt at solving the task.

Watching videos involves coordinating behaviour in real time.

The video recordings showed the facilitator engaging with the same puzzle which the potential learners had just attempted to solve and that they would subsequently manipulate once again. When watching the videos, and as a function of their existing abilities of perception-action, learners might detect the optical patterns specifying the different puzzle pieces and the facilitator's hands, as well as the optical patterns specifying the changes brought about by the manipulations. This may change aspects of their task-related abilities and have consequences for their behaviour when they attempt the task again. Thus, watching the demonstration videos is here seen, not as a matter of acquiring information (*sensu* represented knowledge) from the facilitator, but rather as a matter of sampling and detecting the information (*sensu* meaningful patterns of light and sound) made available by the facilitator, with possible long-lasting (developmental) effects in task-related abilities (see **Figure 7-16**).

In order to explore environmental information, potential learners must coordinate their flow of behaviour and attention to what is happening on the monitor. Here I focused on two aspects of this coordination. On the one hand, I examined the coordination between the time series of the pieces being manipulated by the facilitator (as shown in the videos) and the time series of the pieces being looked at by the potential learners (as captured by the eye tracker). On the other hand, I examined the coordination between the time series of the facilitator looking at the potential learners' eyes (more precisely, at the camera) and the time series of the learners looking at the facilitator's eyes (shown in the monitor). The assumption here is that these periods of behavioural coordination might have persisting effects on the bodily organization of the potential learners, especially in the neuronal networks involved in detecting the patterns which specify task-relevant features of their situation. These effects might become evident later on when participants have another go at the task. For example, learners might perceive relevant aspects which they would otherwise have failed to detect, or they might remember which piece they should pick up given the current configuration of the partially-solved puzzle. Thus, during the observation phase, the facilitator might influence the behavioural flow of the learner

indirectly, i.e. by means of a product of his behaviour (the digital media files). The learner, however, could not affect the behaviour of the facilitator.

Learners in all conditions tended to look at precisely those pieces being manipulated by the facilitator. The study design attempted to influence this attentional coordination by allowing learners to see either the facilitator's face or face blurred, and by allowing learners to either listen to the facilitator's speech or not. Overall, learners looked more at the facilitator's face when it was visible compared to blurred, and even more if the audio was also played. Moreover, this behaviour tended to be triggered in response to the facilitator looking 'at them' (i.e., at the camera) and seemed to play a motivational role since the face is not relevant to solving the task. However, there was no evidence that seeing the facilitator's face, compared to face blurred, affected how leaners coordinated their eye movements with respect to the pieces being manipulated by the facilitator. Together with the overall high proportion of time spent looking at the pieces, this is consistent with the hand-eye coordination route to joint attention suggested by Yu and Smith (2013), rather than the more widely studied gaze following route.

Given that the verbalization produced by a speaker can influence the listener's attention, and given that the facilitator explicitly intended to instruct about the manipulations being performed, I predicted that listening to the facilitator's speech, compared to not, should have a positive overall effect on how leaners coordinated their eye movements with respect to the pieces being manipulated by the facilitator. There was no evidence supporting this. The values of the coordination variables, which indicate how learners moved their eyes while watching the demonstration videos, were very similar across all trials and across conditions.

On the other hand, there was some evidence that listening to the facilitator's speech affected eye movements in one condition. When the facilitator's face was blurred and the audio was played, learners synchronized their eye movements with the facilitator's actions progressively more across iterations, but they stayed synchronized for relatively shorter periods compared to participants in two other conditions (noFACE_noAUDIO and FACE_AUDIO). This suggests that listening to the audio might have prompted learners to explore other parts of the screen (thus reducing the similarity in the cross-recurrences across trials), especially in the initial trials. Thus, it is possible that listening

to the facilitator's speech might influence how learners explore their environments visually even if this was not prominent here. In this case, the dynamics of eye movements was primarily influenced by the optic information created by the manipulation being shown to them, rather than by seeing the facilitator's face or listening to him speak.

Solving puzzles involves coordinating behaviour in real time.

During the test phase, potential learners engaged the task hands-on according to their (changing) abilities to perceive-act and in relation to the (changing) spatial configuration of the puzzle pieces. As a consequence of the forces thus applied on the pieces, their initial spatial arrangement might be transformed in ways we observers may consider the correct or incorrect solution. Thus, attempting to solve the puzzles is here seen, not as a matter of expressing previously acquired information (*sensu* represented knowledge) in overt behaviour, but rather as a matter of sampling and detecting the information (*sensu* meaningful patterns of light) made available by the pieces and using it to guide task-relevant manipulative actions in real time, with possible long-lasting (developmental) effects in task-related abilities (see **Figure 7-16**).

In order to explore environmental information and solve the puzzles, participants must coordinate their flow of behaviour and attention in relation to what is happening with their bodies and the pieces. However, here I did *not* examine behavioural coordination during the test phase. Rather, I coded participants' performance in these tests and interpreted them as an index of their task-related abilities. The baseline test indicated their abilities at the beginning of the learning period. The subsequent tests, following the iterative procedure of the study design, were used to capture possible changes in relevant body parts (especially neuronal networks) as the activity unfolded. I interpreted improvement in performance across trials as indicating, not the acquisition of representations, but systematic changes in task-related body parts in the expected direction – in other words, learning.

Behavioural coordination during demonstration affects subsequent task performance

Participants learned faster when they could not only see the facilitator demonstrate the solutions to the tasks but also listen to him talk about it. In addition, participants learned relatively slower when they could see the facilitator's face compared to face blurred (although the evidence for this negative effect was weak). By manipulating the puzzles

and talking about what he was doing, the facilitator structured the ambient arrays of light and sound, thus creating optic and auditory information that was publicly available (see **Figure 7-16**). The learners might detect and make sense of this information according to their own existing skills, and possibly use it to guide their own actions when they later engage the tasks.

Looking at precisely the pieces being manipulated by the facilitator should help learners perceive the relevant sequence of events that compose the solution to the tasks. Therefore, I predicted that learning should be positively associated with variables measuring the coordination between the pieces being manipulated by the facilitator and the pieces being looked at by the learners. The association between coordination variables and performance provided an intriguing set of results, most of it supporting the prediction. In general, participants who coordinated their eye movements more, compared to less, with the manipulations had a higher probability of success, after we accounted (statistically) for the effects of condition and iteration. Similarly, participants who synchronized their eye movements more and for longer bouts with the facilitator's actions, compared to less and for shorter bouts, also had a higher probability of success. However, this last point was the case only if the facilitator's face was blurred. When his face was clearly visible, on the other hand, the statistical association was reversed: more synchronization hampered learning. It remains unclear how to make sense of this pattern. I provide one alternative below.

When the facilitator's face is blurred, learners can focus on sampling the optic information made available by his actions with the puzzle pieces. In the current context, it seems sensible to suggest that looking at precisely those pieces being manipulated should allow learners to detect relevant aspects of the situation as it evolves in time, and this should be beneficial to learning. When the face is visible, this general reasoning still applies. However, the presence of the face and, especially, the facilitator's behaviour of 'looking at' the learner, changes the context. Consider the evidence. On the one hand, performance tended to be lower when the face was visible compared to blurred. Thus, the presence of the face in the field of view might have distracted the learners somewhat from the demonstration. On the other hand, in those conditions where the facilitator's face was visible, learners who responded by 'looking back' tended to have a higher probability of

success compared to those who did not. I speculated that this behaviour might indicate motivation and commitment to the task.

Thus, it seems that, in the cases where the face is visible, attending too much at the pieces – thus synchronizing more (DET) and for longer (L) – indicates the participants were less engaged. It might be that their eyes are just moving towards the movement specified in the optic array, which is known to capture the attention of the eyes (Mital, Smith, Hill, & Henderson, 2011). Note that this explanation does not make sense in the conditions in which the face was blurred, since in these cases participants only very rarely looked at the blurred blob and this did not correlate with performance. The trend of a positive effect of DET_{face} on success, precisely in the conditions where the face was visible, seem to be consistent with this suggestion.

Learning socially: not transmission, but development

The learning situation produced by the study design involved unidirectional influence from the facilitator to the potential learners. This unidirectional influence, however, should not be confused as indicating a linear transmission of information (*sensu* represented knowledge) from facilitator to learners. Rather, it indicates an asymmetry. On the one hand, the facilitator makes information (*sensu* meaningful patterns of light and sound) available for the potential learners who, in turn, must detect and make sense of it on their own, possibly changing (i.e. developing) their task-solving ability as a result. In contrast, the environmental information made available by the learners' presence and behaviour during this period cannot possibly influence the facilitator.

Although listening to the verbal instructions did facilitate learning, in this case it did not affect how leaners coordinated their eye movements with respect to the pieces being manipulated by the facilitator. Authors working within the currently dominant approach might be tempted to interpret these results as supporting the view of social learning as transmission of information (*sensu* represented knowledge). They might do so by arguing that one of the main roles of speech is to transmit 'knowledge' (e.g., Csibra & Gergely, 2009) and that the behavioural coordination involving mutual responsiveness (one aspect of communication that I have been emphasizing in this thesis) is only marginally important to social learning. I agree that speech may be used to 'communicate knowledge' in the sense that listeners might interpret what they hear in ways consistent

with what the speaker intended to convey. However, this need not imply a commitment to the view of social learning as transmission of information (*sensu* represented knowledge). Although this study was not designed to investigate this in detail, anecdotal evidence suggests a possible explanation for speech having a positive effect on learning and virtually no effect on coordination, in line with the current view. During debriefing, one participant spontaneously suggested that she tried to memorize what the facilitator said and later, when she was engaging with the task, she would follow the instructions as she remembered them. This suggests a plausible hypothesis for the role of memory in cases of social learning involving speech worth examining in further studies. Consider that the narrative structure of the verbal instructions correlates in real time with the narrative structure of the demonstration itself (see Appendix). Both imply the same history of transformations in which the initial situation (puzzle pieces scattered on the desk) becomes, through a flow of manipulative behaviours, the final situation (puzzle pieces forming the desired structure). The verbal instructions provide learners with a description of which pieces are relevant at some period and which actions they are involved in at each step. The suggestion is that participants might later re-enact this sequence of instructions as they remember it and use it to guide their own attention and behaviour in real time. Remembering is here understood, not in terms of retrieving data structures stored in a computer-like memory, but in terms of how (auditory) experience can have lasting (developmental) effects in the individual's ability to simulate similar (auditory) experience in the future. The suggestion is that this simulating auditory experience involves enacting dynamical patterns of neurophysiological activity that may constrain (reduce the degrees of freedom of) other components of the ongoing global neurophysiological activity and, therefore, the individual's behavioural flow. Remembering (simulating auditory perception of) the verbal instructions might explain the positive effect of audio on learning in this case without systematic differences in eye movements across conditions. Future work should investigate the roles remembering, planning, and imagining, play across different cases of learning socially in ways consistent with a radical embodiment perspective.

Characterizing social influences on learning in the current case

Watching video recordings of the facilitator performing demonstrations gave participants the opportunity to embody implicit knowledge about solving a set of puzzles. In Chapter

5, I suggested three criteria to clarify the scope of any study of social learning. With regard to the time scale, here I focused on learning that might occur within the short time scale of several minutes. With regard to the spatial scale, the learning situation unfolded within a lab-based, controlled environment. The main components under analysis were the learners, the task and supporting materials, and the stimuli playing the role of 'social influence'. With regard to the learning outcome, this study focused on skill learning by using a set of construction puzzles as the target task.

I also suggested five criteria to clarify the 'social influence' under scrutiny. With regard to the general role played by the facilitator in the learning process, here the facilitator engaged in the same task in which the learner was judged, thus acting as a demonstrator. Only the facilitator could influence the learner, thus the influence was unidirectional (from facilitator to learner). The learner did not interact with the facilitator in real time but rather with recorded videos, thus the influence was indirect. However, given the use of recorded videos, the stimuli created flowing optic arrays similar to what would be available if the learners watched the demonstrations live. With regard to their prior intentions, the facilitator explicitly intended his behaviour to promote learning, and the learners explicitly committed to attempt learning the task. In conclusion, this study focused on one example of how learning can be influenced by the presence and behaviour of others around us. From the perspective of radical embodiment, this process is understood in terms of the facilitator influencing how learners engage with the task at hand and, in so doing, how learners might transform their existing skills. This is an alternative to saying that the facilitator transmits knowledge or representations to them. Focusing on coordination explicitly and quantitatively might improve our understanding of this broad class of phenomena. This study provided some conceptual, methodological and analytical tools in this direction.

Chapter 8 General Discussion

This thesis set out to outline a novel way of thinking about social learning that was motivated by a set of critical developments across relevant disciplines. The previous chapters explored ideas from developmental systems theory, radical embodied cognitive science, as well as process and relational thinking, and also reported empirical studies examining behavioural/attentional coordination and its relation to social learning. In this final chapter, I will first provide a summary of findings. I will then discuss why the metaphor of transmission of information fails to capture important aspects of the phenomena of interest, and provide an alternative. I will then suggest areas for future work before concluding.

8.1 Summary of findings

References to 'information' are ubiquitous in the literature on social learning, usually in three variations: genetic information, asocial information, and social or cultural information. Therefore, I attempted to clarify the different meanings that the term might acquire, especially in the context of development and behaviour (Chapters 2 to 5). Crucially, the statement that 'situation A might carry *information about* situation B' was understood in terms of lawful or normative constraints that hold between A and B (Barwise & Perry, 1981; Devlin, 2006; Dretske, 1981; Israel & Perry, 1990). However, in this sense, there was no ground to justify the view that genetic information corresponds to "a set of instructions" or "a recipe" for how to make the organism (Maynard Smith, 2000, p. 187). On the other hand, reference to 'information' in the context of behaviour and social learning commonly indicates a representational-computational view of cognition. Here, the term is used colloquially in the sense of mental representations, considered to be analogous to computer data structures (Thagard, 2014).

To avoid the nature-culture distinction, along with the associated distinctions nature-nurture, biology-culture, body-mind, individual-society, I attempted an original integration of developmental systems theory, radical embodied cognitive science, and relational thinking. The overall strategy consisted in conceiving of biological and behavioural form as being continually produced within a developmental system, considered as a network of causal relations among processes unfolding in time. Depending on the focus of interest, this network of causal influences might be provisionally dissected in different ways, giving prominence to some components or processes while leaving others in the background. In Chapters 6 and 7, for example, I focused on the learner, the facilitator, and the materials and objects involved in the respective activity, as composing the developmental system under scrutiny.

Chapters 3 focused on ontogeny – the entire lifespan of an organism – conceived of as a historical (dynamical), relational, constructive and contingent process in which a zygote or a corresponding initial configuration persists under transformation until it eventually dies. Chapter 4 focused on behaviour as one aspect of the flow of ontogeny. Behaviour was understood, not in terms of computation with representations, but in terms of the dynamic pattern produced by the entire animal-environment system. The related notions of 'environmental information', 'affordance', and 'ability' were used to make sense of how animals can perceive and act upon their environment. Environmental information refers to meaningful patterns that appear in the ambient arrays of light, sound, chemicals, etc. (J. J. Gibson, 1979/2015). These patterns are meaningful in the sense that they are linked to other aspects of the world by lawful or normative constraints and can therefore inform the skilful animal about them.

Some of the information available in the environmental situation is especially relevant for behaviour because it specifies opportunities for action or affordances. Affordances refer to what abilities animals might exercise, given the situation in which they find themselves. Abilities refer to what animals might perceive and do, given their bodily organization and the affordances currently present. I accepted the suggestion by Chemero (2009) to view abilities and affordances dynamically rather than as properties of the environment (Turvey et al., 1981), and I argued for the need to include intentions in the picture. Instead of saying that behaviour is the expression of information (*sensu*

represented knowledge) stored and processed in the nervous system, I suggested that the flow of behaviour, i.e. the lived sequence of perception-action events, results from the dynamic, real-time interplay of (1) the current (changing) affordances of the environment; (2) the current (changing) abilities of the animal; (3) the (more or less specific) intentions of the animal (or, alternatively, differences in the valence associated with aspects of the environment); and (4) random or unpredictable fluctuations. Learning was then defined in terms of developmental changes in the body parts involved in abilities, especially, but not restricted to, relevant neuromuscular networks and the coordination dynamics they may produce.

Chapter 5 focused on living and learning socially. Animals commonly live within multispecies communities. Therefore, the environmental information created by the presence and behaviour of one animal is publicly available and might be detected by other, skilful animals. Thus, animals may influence each other's behaviour as they go about living within a common environment. Moreover, because animals might be mutually sensitive to each other's presence and behaviour, they might become mechanically or informationally coupled for some period of time. If the behaviour of one animal is systematically influenced by another animal, this might have developmental consequences for that animal's abilities. These cases are captured by the broad notion of social learning as "learning that is facilitated by observation of, or interaction with, another individual (or its products)" (Hoppitt & Laland, 2013, p. 4). Thus, consistent with a view of ontogeny as a historical (dynamical), relational, constructive and contingent process, I argued that learning is not about acquiring 'information' or representations relevant to the target task, and social learning is not about acquiring 'cultural information.' Instead, learning, and learning socially, are here seen as possible material consequences of relational developmental processes. More specifically, learning denotes a change in relevant body parts (e.g. bones, joints, muscles, neuronal networks) underlying some ability of interest. We observers might infer that such changes have occurred by comparing the individual's performance in some test before and after the 'learning' period under consideration. I suggest this (meta)theoretical orientation allows us to investigate how living together may influence the flow of behaviour and the development of abilities without committing to distinguishing 'culture' as a source of representations or as a domain fundamentally distinct from 'biology'.

Chapters 6 and 7 corresponded to the empirical part of this project, which provided case studies used to apply the way of thinking about social learning developed in the previous chapters. In Chapter 6, I examined a joint making activity and explored analytical tools to capture the dynamical constitution of coupled attention and its association with learning. In Chapter 7, the task was defined in such a way as to allow for better control and a finer spatiotemporal resolution, however at the cost of losing the bidirectional influence that depends on the real-time interaction between individuals. I used these studies as opportunities to practice an alternative way of thinking and writing about social learning consistent with the (meta)theoretical framework put forward in the previous chapters. The phenomena of interest were described and explained, not in terms of genetic or cultural information but in terms of ecological information and the transformation of abilities consequent on experience. The influence observed from facilitator to learner was described, not in terms of the transmission of information (*sensu* represented knowledge) but in terms of the creation of environmental information (*sensu* meaningful patterns of light and sound), the coordination of the learner's behaviour in relation to this information, and the association between this coordination and learning. Both studies provided quantitative evidence that behavioural coordination is associated with learning.

8.2 Social learning: not transmission, but development

The metaphor of 'transmission of information' is ubiquitous in the dominant literature on social learning and using it might indeed simplify the verbal description of what is going on in a given case. Some authors might find this strategy especially useful when examining larger spatiotemporal scales such as the 'passing on' of knowledge across several generations rather than between two individuals. Transmission talk might also sound reasonable in studies of large-scale populational processes using computational algorithms in which the target behaviour is modelled by variables involved in iterative 'copy-paste' procedures. However, I have been arguing against the transmission metaphor throughout this thesis. **Figure 8-1** shows two schematic representations of social learning conceived of as a process of transmission and as a process of relational development. In addition, **Figure 8-2** shows additional aspects of the common environment in which the facilitator and the learner live together. This is intended to highlight the broader material context in which social learning may occur.

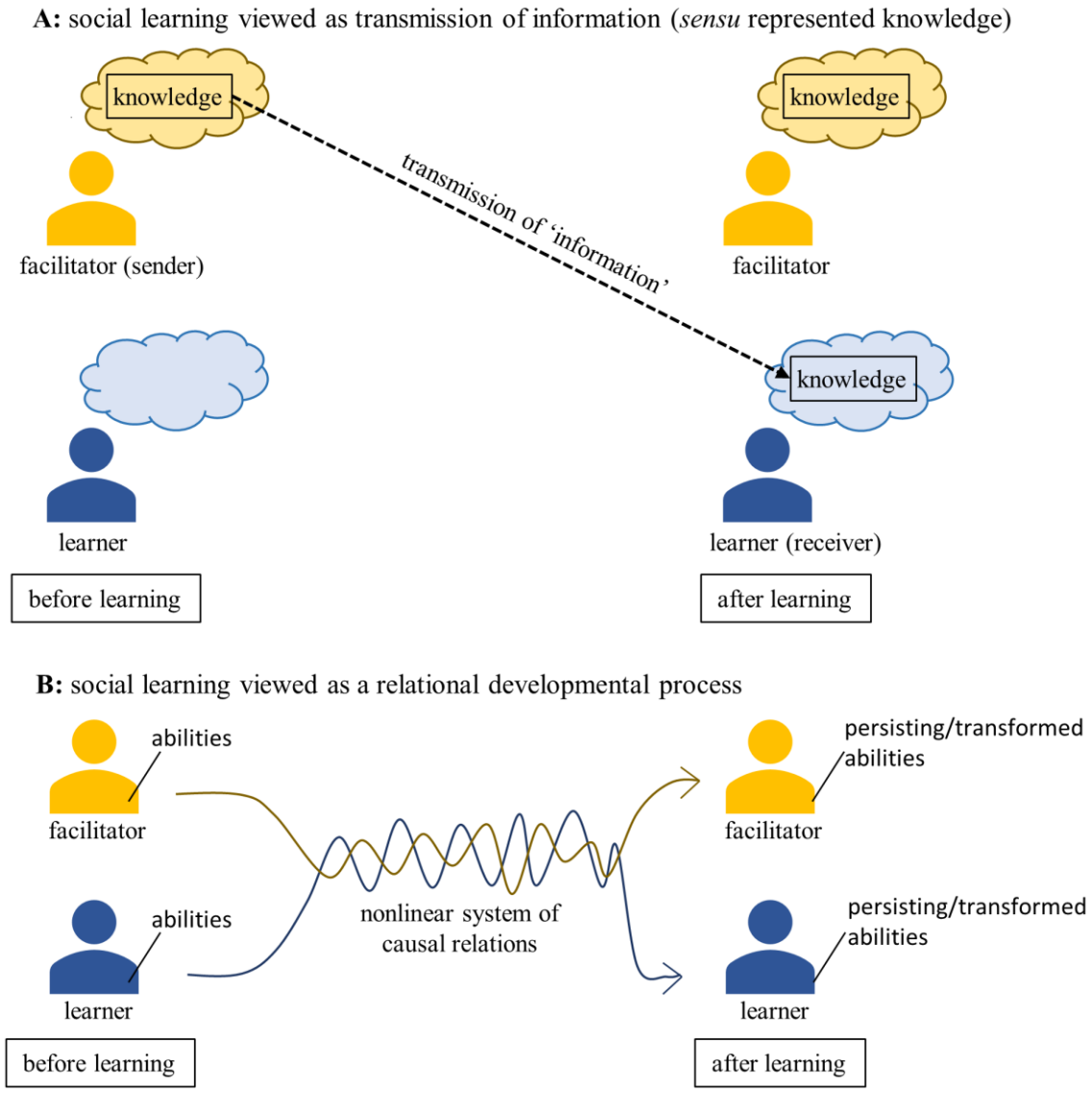


Figure 8-1. Schematic representation of social learning conceived of as a process of transmission of 'information' (A) and as a relational developmental process (B).

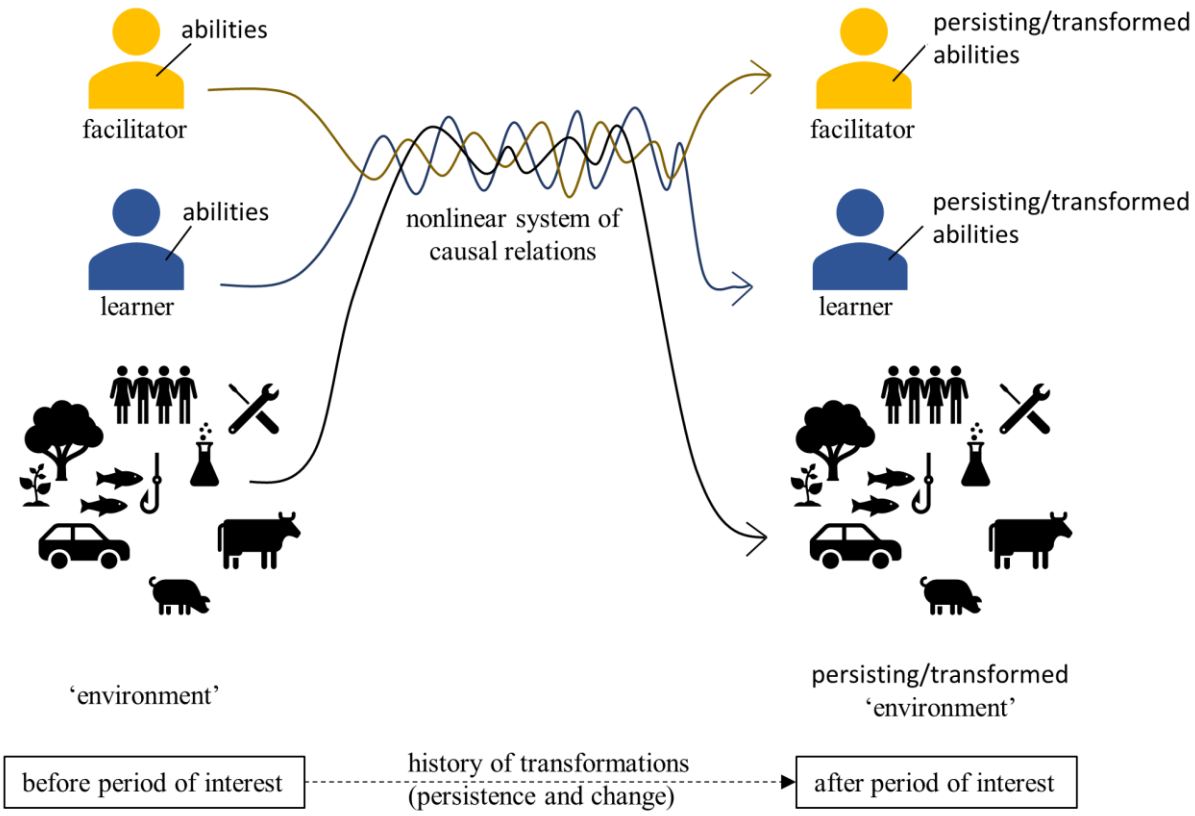


Figure 8-2. Schematic representation of social learning conceived of as a relational developmental process, with additional aspects of the facilitator's and learner's common environment explicitly indicated.

I want to highlight four major problems with the transmission metaphor. The first problem is that transmission talk is commonly associated with a representational view of the mind, which is rejected here. From the perspective being put forward, knowledge is not conceived of as being contained in mental representations. Rather, knowledge is conceived in relational and dynamical terms, as being supported by the material components of the individual's body (e.g., bones, joints, neuromuscular networks) and the temporally extended relations they may enter with task-relevant aspects of their environment as a function of their biomechanical and coordinative properties. The bodily components associated with knowledge cannot be 'transmitted' from one individual to another, since each individual must develop their own body parts. Thus, I find it difficult to make sense of theoretical statements such as that novel actions can, in some cases of social learning, be "acquired by the observer directly through observation" (Hoppitt & Laland, 2013, p. 73).

The transmission metaphor stems from the mathematical theory of communication, which involves a notion of information as reduction of uncertainty (Shannon, 1948a, 1948b). I suggest it is possible to use the term information consistently, in the context of social learning, even when avoiding the transmission metaphor. This can be done by using the notion of ecological information (J. J. Gibson, 1979/2015). Thus, rather than saying that the facilitator transmits information (*sensu* represented knowledge) to the learner, we might say (1) that the presence and behaviour of the facilitator creates environmental information (*sensu* meaningful patterns in the environment) that becomes publicly available and (2) that the learner, constrained by his or her current abilities, might pick up and use this environmental information to guide the flow of behaviour. As a consequence of this relational process, some aspects of the organization of the learner's body persist while other aspects might change systematically affecting his or her abilities as we observers detect them. The transmission metaphor does not capture this non-representational notion of knowledge and how it can be 'passed on' as a consequence of a relational history.

A similar argument involves communication. I agree that humans may use speech, and other inscriptive practices such as written text, diagrams, and artefacts, to communicate what they know about the world. However, in communication (verbal, written, pictorial, and so on), the facilitator is not 'transmitting' knowledge in any material sense, but rather creating environmental information in the sense of patterns of light, sound, surfaces, chemicals, and so on. Potential learners may detect some of these patterns. When detecting these patterns, rather than having knowledge transmitted to them, ready to use, potential learners must make sense of what they see, hear, touch, etc., according to their own abilities to perceive, act, imagine, remember, plan, and so on, that they developed thus far.

The second major problem with the transmission metaphor is that it neglects mutual sensitivity in facilitating-learning relations. In cases involving bidirectional influence, the target learner may detect and respond to the behaviour of the facilitator, and the facilitator may detect and respond to the behaviour of the learner, as they keep on living together. As a result of their historical interaction, their bodies might change systematically. This mutual sensitivity brings about a history of nonlinear causal influences not captured by

the idea of linear transmission. In their influential work on teaching in nonhuman animals, Caro and Hauser (1992) acknowledge that facilitators might behave differently in the presence or absence of the learners. However, they explicitly chose to neglect the issue of sensitivity because they believe that, "from an evolutionary perspective that focuses on the effects of teaching, there is no a priori reason why one should necessarily expect more or less sensitive mechanisms to have a greater or lesser effect on the reproduction and survival of the pupil" (Caro & Hauser, 1992, p. 155). Here I chose not to neglect the importance of sensitivity and examined it in terms of bidirectional coordination established and sustained throughout the activity. The empirical study reported in chapter 6 explored ways to capture bidirectional coordination for quantitative analysis and provided evidence of statistical association between coordination measures and learning outcome.

The third major problem with this metaphor is that it supports an artificial distinction between teaching-related behaviour ('sending the information') and learning-related behaviour ('receiving the information') as logically independent or "orthogonal to each other" (Hoppitt et al., 2008). Here I offered an alternative account of the relation between teaching and social learning. More specifically, teaching-related behaviour and learning-related behaviour are seen as complementary aspects of a single history of co-regulation of behaviour involving the facilitator and the learner. On the one hand, the facilitator is sensitive to what the learner is doing and behaves with the (more or less specific) intention to teach. On the other hand, the learner is influenced by what the facilitator is doing and might change as a result. Thus, teaching-learning relations form a subset of the broader class of facilitating-learning relations.

The fourth major problem with the transmission metaphor is that it fails to capture the causal roles played by tools and materials, such as in learning through making together. Through their behaviour, animals are able to alter their environment by making things and may indeed be able to imagine the desired form and plan their course of action more or less precisely, according to their previous experience. Imagining and planning are here understood in dynamical terms, as simulated experience of perception-action that may constrain the ongoing activity of neuromuscular networks, including those more directly associated with the occurring overt behaviour. On the one hand, imagining,

planning, perceiving, and acting, play causal roles in the transformation of materials-becoming-artefact. On the other hand, in a making activity, the tools and materials effectively become part of the maker's developmental system, thus influencing the unfolding history of causal relations according to their physicochemical properties. Therefore, making can be conceived of as unfolding within a relational, developmental process in which the makers, the tools, and the materials, affect each other in a seamless history of material, causal relations. As a result of these causal relations, the materials might change in their composition, physicochemical properties, and spatial arrangement, and become new things. The bodies of the makers might also change in their composition, biomechanical and coordinative properties, and spatial arrangement, and support new abilities. Making, therefore, is a form-generating process, and this applies to the artefact as well as to the makers. Learning through making, as a history of transformation involving both persistence and change in materials and in body parts, can thus be seen as nested within the individual's ontogeny.

In sum, what links the facilitator and the learner is not a communication channel transmitting signals carrying some amount of information. Rather, the unskilled individual *becomes* the skilled individual in a history of persistence and change of the individual's entire developmental system. This developmental process might involve uni- or bidirectional influences relative to the facilitator as well as the (changing) material surroundings.

8.2.1 Categories of social learning as genres of developmental histories

Social learning refers to a historical achievement of a set of mutually constraining processes, including the facilitator, the learner, and relevant aspects of their common environment. Therefore, similar to how sets of literary or film works can be categorized within different literary or film genres, I suggest categories of social learning might be best understood as denoting *genres of developmental narratives* rather than mechanisms of transmission of 'information.' In other words, a category of social learning refers to a set of ontogenetic trajectories that share family resemblance according to some criteria. What these criteria should be is a matter of theoretical debate. In Chapter 5, I suggested these criteria might include the time scale, the spatial scale, and the learning outcome of interest; the role of the environmental information made available by the facilitator;

whether the facilitator performs the target task or does something else; whether the influence between facilitator and learner is unidirectional or bidirectional; whether the influence is direct (by means of real-time behavioural coupling) or indirect (by means of persisting changes in the environment); and the intentions of those involved. Further work could examine this suggestion in more detail, e.g. with the intention to produce a principled classification scheme. This might improve our theoretical understanding of social learning processes because of the problems with the current typologies indicated in Chapter 5.

8.3 Further work

8.3.1 Theoretical questions

Further work should focus on advancing the theoretical integration initiated here. Throughout the thesis, I indicated work by social anthropologists produced from a relational perspective. It will be important to engage in further dialogue with this literature more explicitly as each might have insights to offer to the other.

One of the challenges commonly made to those advocating for radical embodied cognitive science is to provide a more robust account of so-called higher cognitive abilities such as intending, planning, remembering, or imagining. It is commonly assumed that it is difficult, if not impossible, for animals to do these things without representing the world around them and manipulating these representations in computational algorithms. For those interested in advancing radical embodied cognitive science, it is of central importance to explore ways to think and talk about the processes indicated by terms such as intending, planning, remembering, and imagining, which have traditionally been regarded as involving mental representations. For example, rather than viewing memory as the retrieval of represented knowledge, I suggest we focus on planning and remembering as activities in which the individual simulates perceptual experience. This is consistent with Gibson's hypothesis that "a perceptual system that has become sensitized to certain invariants and can extract them from the stimulus flux can also operate without the constraints of the stimulus flux" (J. J. Gibson, 1979/2015, p. 244). There is currently exciting work being produced in this direction (Bruineberg, Chemero, & Rietveld, 2018; Juarrero, 1999; Kiverstein & Rietveld, 2018). The context of social

learning provides excellent opportunities to explore activities involving 'higher' cognition from the current perspective.

8.3.2 Methodological questions

It is reasonably straightforward to use cross recurrence quantification analysis to examine the coordination between two time series (e.g., one representing a facilitator and another the target learner). However, oftentimes learning involves situations in which larger groups of people (and other animals) interact in real time. It is less straightforward to envision how to apply this methodology to such cases. Thus, it would also be important to invest in further work focusing on advancing relevant methodology that might be used to study behavioural coordination in larger groups.

In chapter 4, I explicitly included the intentions of the individual as one process causally involved in producing the flow of behaviour. In chapter 5, I argued that future work should focus on proposing operational criteria to identify the different intentions constraining the behaviour of facilitators and learners in specific cases of social learning. With this in mind, I suggest future work might focus on proposing and testing methodology that might be used to infer the intentions associated with teaching- and learning-related behaviour. This would be particularly helpful to those interested in conducting comparative studies of teaching-learning relations.

8.3.3 Empirical questions

Further empirical work might use cross recurrence quantification analysis (or other analytical tools) to systematically examine variation in coordination variables produced by sets of dyads, or larger groups if possible, engaging together in what might be described as 'the same' (or relatively similar) learning contexts. For example, we might examine how learners from different age groups, and/or from different historical backgrounds, coordinate their behaviour during a situation we might call 'intentional demonstration' (e.g., similar to the subtask TOOL_DEMO in the joint making activity reported in Chapter 6) or 'assisted practice' (e.g. similar to the subtask TOOL_USE in the same study). Additionally, we might examine the association between the coordination produced during such periods and measures of performance in a subsequent 'practice' period.

In addition, we might investigate whether participants engaging in different learning situations produce systematically different patterns of coordination. From a comparative perspective, we might investigate how individuals from different species (say, humans and other primates) coordinate their behaviour in what might be described as 'the same' (or a relatively similar) learning context such as 'intentional demonstration,' 'incidental (unintentional) demonstration,' or 'assisted practice'. If different learning categories do indeed produce systematically different signatures of coordination, this knowledge might later be used to provide evidence about what might be going on in novel cases.

8.4 Conclusion

This project was motivated by a dissatisfaction with the dominant view that "we are largely what our genes and our culture make us" (Richerson & Boyd, 2005, p. 6), notwithstanding however intricately these factors might be said to 'interact' to produce the individual person or animal. The alternative endorsed here is to think of form – including the form of the organism, the form of individual behaviour, and the form of group-level patterns of behaviour – as persisting under historical transformations rather than being imposed, expressed, or transmitted, and to investigate the form-generating processes involved in the appearance and historical transformation of the phenomena of interest.

The distinction between a domain of nature and a domain of culture or society has been under criticism within the social sciences for decades (Descola & Pálsson, 1996; Geertz, 1973; Ingold, 2000, 2011a; Ingold & Pálsson, 2013; Latour, 1993; Lévi-Strauss, 1966; Stengers, 2000; Toren, 2012). This criticism is directed to a set of so-called Modern oppositions including biology-culture, body-mind, and individual-society. Similarly, the distinction between a domain of nature (biology, genes, innate, instinct) and a domain of nurture (culture, environment, acquired, learned) has also been under criticism within the biological sciences for decades (Gottlieb, 1976, 1997, 2007; Keller, 2010; Kuo, 1967; Lehrman, 1953, 1970; Lewontin, 1983; Oyama, 1985/2000, 2000b, 2015; Oyama et al., 2001). Within the cognitive sciences, the distinctions between the body and the mind, and between the organism and the environment have also been under criticism for decades (Chemero, 2009; E. J. Gibson & Pick, 2000; J. J. Gibson, 1966, 1979/2015; Kelso, 1995;

Thelen & Smith, 1994; Thompson, 2007; Van Gelder, 1995; Varela et al., 1991).In this thesis, I explored what difference it might make to accept these criticisms and take the positive alternatives they suggest as a starting point to the study of social learning.

In conclusion, this thesis contributes to the field of social learning by suggesting a new way of thinking about the relations among ontogeny, cognition, behaviour, the behavioural coordination enabled by living socially, and social learning. This way of thinking is consistent with critical theoretical and philosophical developments that are currently neglected in the field. This thesis also endorses and contributes to the field of radical embodied cognitive science by discussing theoretical connections with other areas of research, critically discussing the metaphor of transmission of information, clarifying terminology (such as the use of the term ‘information’), and suggesting areas for future research. While many points raised here must be further clarified and their implications more fully discussed, the present perspective suggests an exciting direction that might improve our understanding of how we live and learn socially.

Reference List

Alberts, B., Wilson, J. H., & Hunt, T. (2008). *Molecular biology of the cell* (5th ed.). New York: Garland Science.

Allen, J., Weinrich, M., Hoppitt, W. J. E., & Rendell, L. (2013). Network-Based Diffusion Analysis Reveals Cultural Transmission of Lobtail Feeding in Humpback Whales. *Science, 340*(6131), 485-488. doi:10.1126/science.1231976

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419-439. doi:10.1006/jmla.1997.2558

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., & Sheldon, B. C. (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature, 518*(7540), 538-541. doi:10.1038/nature13998

Barwise, J. (1983). Information and Semantics. *Behavioral and Brain Sciences, 6*(1), 65-65.

Barwise, J., & Perry, J. (1981). Situations and Attitudes. *Journal of Philosophy, 78*(11), 668-691. doi:10.2307/2026578

Barwise, J., & Seligman, J. (1997). *Information flow : the logic of distributed systems*. Cambridge ; New York: Cambridge University Press.

Bateson, G. (1972). *Steps to an ecology of mind : collected essays in anthropology, psychiatry, evolution, and epistemology*. San Francisco: Chandler Pub. Co.

Boesch, C. (1991). Teaching among wild chimpanzees. *Animal Behaviour, 41*(3), 530-532. doi:10.1016/S0003-3472(05)80857-7

Boesch, C., & Boesch, H. (1981). Sex differences in the use of natural hammers by wild chimpanzees: A preliminary report. *Journal of Human Evolution, 10*(7), 585-593. doi:10.1016/S0047-2484(81)80049-8

Boesch, C., & Boesch, H. (1984). Possible causes of sex differences in the use of natural hammers by wild chimpanzees. *Journal of Human Evolution, 13*(5), 415-440. doi:10.1016/S0047-2484(84)80055-X

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Taï National Park. *American Journal of Physical Anthropology, 78*(4), 547-573. doi:10.1002/ajpa.1330780410

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Brentano, F. (1995). *Psychology from an empirical standpoint* (O. Kraus & L. L. McAlister Eds. Paperback ed.). London; New York: Routledge.

Briggs, J. L. (2002). Language dead or alive: What's in a dictionary? *Topics in artic social sciences, 4*, 69-82.

Bruineberg, J., Chemero, A., & Rietveld, E. (2018). General ecological information supports engagement with affordances for 'higher' cognition. *Synthese*. doi:10.1007/s11229-018-1716-9

Camazine, S. (2001). *Self-organization in biological systems*. Princeton, N.J.: Princeton University Press.

Caro, T. M., & Hauser, M. D. (1992). Is there teaching in nonhuman animals? *Quarterly Review of Biology, 67*(2), 151-174. doi:10.1086/417553

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2017). Stan: A Probabilistic Programming Language. *2017, 76*(1), 32. doi:10.18637/jss.v076.i01

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4).

Carpenter, M., Tomasello, M., & Savage-Rumbaugh, S. (1995). Joint Attention and Imitative Learning in Children, Chimpanzees, and Enculturated Chimpanzees. *Social Development, 4*(3), 217-237. doi:10.1111/j.1467-9507.1995.tb00063.x

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton, N.J.: Princeton University Press.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, Mass.: MIT Press.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*(1), 62-81. doi:10.1016/j.jml.2003.08.004

Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: an R package. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.00510

Coco, M. I., Dale, R., & Keller, F. (2018). Performance in a Collaborative Search Task: The Role of Feedback and Alignment. *Topics in Cognitive Science, 10*(1), 55-79. doi:10.1111/tops.12300

Corp, N., & Byrne, R. W. (2002). The ontogeny of manual skill in wild chimpanzees: Evidence from feeding on the fruit of Saba florida. *Behaviour, 139*, 137-168. doi:10.1163/15685390252902328

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153. doi:10.1016/j.tics.2009.01.005

d'Errico, F., & Stringer, C. B. (2011). Evolution, revolution or saltation scenario for the emergence of modern cultures? *Philosophical Transactions of the Royal Society B-Biological Sciences, 366*(1567), 1060-1069. doi:10.1098/rstb.2010.0340

Dale, R., Fusaroli, R., Duran, N. D., & Richardson, D. C. (2013). The Self-Organization of Human Interaction. In B. H. Ross (Ed.), *Psychology of Learning and Motivation, Vol 59* (Vol. 59, pp. 43-95).

Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in Psychology, 2*. doi:10.3389/fpsyg.2011.00355

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences, 6*(4), 485-507. doi:10.1007/s11097-007-9076-9

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn Sci, 14*(10), 441-447. doi:10.1016/j.tics.2010.06.009

De Jaegher, H., & Froese, T. (2009). On the Role of Social Interaction in Individual Agency. *Adaptive Behavior, 17*(5), 444-460. doi:10.1177/1059712309343822

De Resende, B. D., Ottoni, E. B., & Fragaszy, D. M. (2008). Ontogeny of manipulative behavior and nut-cracking in young tufted capuchin monkeys (Cebus apella): a Perception–action perspective. *Developmental Science, 11*(6), 828-840. doi:10.1111/j.1467-7687.2008.00731.x

Dennett, D. C. (1987). *The intentional stance*. Cambridge, Mass.: MIT Press.

Descola, P., & Pálsson, G. (1996). *Nature and Society: Anthropological Perspectives*: Routledge.

Devlin, K. (2006). Situation Theory and Situation Semantics. *Logic and the Modalities in the Twentieth Century, Vol 7, 7*, 601-664.

Di Paolo, E., & Thompson, E. (2014). The Enactive Approach. *Routledge Handbook of Embodied Cognition*, 68-78.

DiFrisco, J. (2017). Time Scales and Levels of Organization. *Erkenntnis, 82*(4), 795-818.

DiFrisco, J. (in press). Biological Processes: Criteria of Identity and Persistence. In J. Dupre & D. Nicholson (Eds.), *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press.

Dretske, F. I. (1981). *Knowledge and the flow of information* (1st MIT Press ed.). Cambridge, Mass.: MIT Press.

Drummond, H. (1981). The nature and description of behavior patterns. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in Ethology* (Vol. 4, pp. 1-34). New York and London: Plenum Press.

Dupre, J., & Nicholson, D. (Eds.). (in press). *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press.

Eckmann, J. P., Kamphorst, S. O., & Ruelle, D. (1987). Recurrence Plots of Dynamical Systems. *EPL (Europhysics Letters), 4*(9), 973.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften, 58*(10), 465-523.

Feldman, M. W., & Laland, K. N. (1996). Gene-culture coevolutionary theory. *Trends in Ecology & Evolution, 11*(11), 453-457. doi:10.1016/0169-5347(96)10052-5

Fodor, J. A. (1975). *The language of thought*. New York: Crowell.

Fogel, A. (1993). *Developing through relationships : origins of communication, self, and culture*. Chicago: University of Chicago Press.

Fragaszy, D. M., Biro, D., Eshchar, Y., Humle, T., Izar, P., Resende, B., & Visalberghi, E. (2013). The fourth dimension of tool use: temporally enduring artefacts aid primates learning to use tools. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1630). doi:10.1098/rstb.2012.0410

Fragaszy, D. M., Eshchar, Y., Visalberghi, E., Resende, B., Laity, K., & Izar, P. (2017). Synchronized practice helps bearded capuchin monkeys learn to extend attention while learning a tradition. *Proceedings of the National Academy of Sciences of the United States of America, 114*(30), 7798-7805. doi:10.1073/pnas.1621071114

Fragaszy, D. M., & Perry, S. (2003). *The biology of traditions : models and evidence*. Cambridge, UK; New York: Cambridge University Press.

Franks, N. R., & Richardson, T. (2006). Teaching in tandem-running ants. *Nature, 439*(7073), 153-153. doi:10.1038/439153a

Freeberg, T. M. (2000). Culture and courtship in vertebrates: a review of social learning and transmission of courtship systems and mating patterns. *Behavioural Processes, 51*(1-3), 177-192. doi:10.1016/S0376-6357(00)00127-3

Futuyma, D. J. (1998). *Evolutionary biology* (3rd ed.). Sunderland, Mass.: Sinauer Associates.

Galef, B. G. (1988). Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. In Z. T. R. & G. B. G. (Eds.), *Social Learning: Psychological and Biological Perspectives* (pp. 3-28). Hillsdale, NJ: Erlbaum.

Gánti, T. (2003). *The principles of life*. Oxford ; New York: Oxford University Press.

Geertz, C. (1973). *The interpretation of cultures : selected essays*. New York,: Basic Books.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (Third edition. ed.). Boca Raton: CRC Press.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge ; New York: Cambridge University Press.

Gerstman, B. S., & Chapagain, P. P. (2005). Self-organization in protein folding and the hydrophobic interaction. *Journal of Chemical Physics, 123*(5).

Gibson, E. J. (1988). Exploratory-Behavior in the Development of Perceiving, Acting, and the Acquiring of Knowledge .9. *Annual Review of Psychology, 39*, 1-41. doi:10.1146/annurev.ps.39.020188.000245

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. Oxford ; New York: Oxford University Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston,: Houghton Mifflin.

Gibson, J. J. (1979/2015). *The ecological approach to visual perception* (Classic Edition ed.). New York and London: Psychology Press.

Gilbert, S. F. (2000). *Developmental biology* (6th ed.). Sunderland, Mass.: Sinauer Associates.

Gilbert, S. F. (2010). *Developmental biology* (9th ed.). Sunderland, Mass.: Sinauer Associates.

Godfrey-Smith, P. (2004). Genes do not Encode Information for Phenotypic Traits. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science* (pp. 275-289). Malden, MA: Blackwell Publishing.

Godfrey-Smith, P. (2007). Information in Biology. *Cambridge Companion to the Philosophy of Biology*, 103-119.

Goodall, J. (1964). Tool-Using and Aimed Throwing in a Community of Free-Living Chimpanzees. *Nature, 201*, 1264. doi:10.1038/2011264a0

Gottlieb, G. (1976). Conceptions of Prenatal Development - Behavioral Embryology. *Psychological Review, 83*(3), 215-234. doi:10.1037/0033-295x.83.3.215

Gottlieb, G. (1997). *Synthesizing nature-nurture : prenatal roots of instinctive behavior*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science, 10*(1), 1-11. doi:10.1111/j.1467-7687.2007.00556.x

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462-466. doi:10.1111/1467-9280.02454

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*(4), 274-279. doi:10.1111/1467-9280.00255

Griffiths, A. J. F. (2008). *Introduction to genetic analysis* (9th ed.). New York: W.H. Freeman and Co.
Griffiths, P. E. (2001). Genetic information: A metaphor in search of a theory. *Philosophy of Science, 68*(3), 394-412. doi:10.1086/392891
Griffiths, P. E., & Gray, R. D. (1994). Developmental Systems and Evolutionary Explanation. *Journal of Philosophy, 91*(6), 277-304. doi:10.2307/2940982
Haken, H. (1977). *Synergetics : an introduction : nonequilibrium phase transitions and self-organization in physics, chemistry, and biology ; with 724 figures*. Berlin ; New York: Springer-Verlag.
Haken, H. (1983). *Synergetics : an introduction : nonequilibrium phase transitions and self-organization in physics, chemistry, and biology* (3rd rev. and enl. ed.). Berlin ; New York: Springer.
Haken, H. (1996). *Principles of brain functioning : a synergetic approach to brain activity, behavior, and cognition*. Berlin ; New York: Springer.
Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*(5), 347-356. doi:10.1007/bf00336922
Helfman, G. S., & Schultz, E. T. (1984). Social Transmission of Behavioral Traditions in a Coral-Reef Fish. *Animal Behaviour, 32*(May), 379-384. doi:10.1016/S0003-3472(84)80272-9
Hernandez-Hernandez, V., Niklas, K. J., Newman, S. A., & Benitez, M. (2012). Dynamical patterning modules in plant development and evolution. *International Journal of Developmental Biology, 56*(9), 661-674. doi:10.1387/ijdb.120027mb
Heyes, C. (1994). Social-Learning in Animals - Categories and Mechanisms. *Biological Reviews of the Cambridge Philosophical Society, 69*(2), 207-231. doi:10.1111/j.1469-185X.1994.tb01506.x
Heyes, C. (2012). What's social about social learning? *J Comp Psychol, 126*(2), 193-202. doi:10.1037/a0025180
Heyes, C. (2016). Blackboxing: social learning strategies and cultural evolution. *Philos Trans R Soc Lond B Biol Sci, 371*(1693). doi:10.1098/rstb.2015.0369
Hobaiter, C., Poisot, T., Zuberbuhler, K., Hoppitt, W. J. E., & Gruber, T. (2014). Social Network Analysis Shows Direct Evidence for Social Transmission of Tool Use in Wild Chimpanzees. *Plos Biology, 12*(9). doi:10.1371/journal.pbio.1001960
Hogan, J. A. (2017). *The Study of Behavior: Organization, Methods, and Principles*. Cambridge: Cambridge University Press.
Hohmann, G., & Fruth, B. (2003). Culture in Bonobos? Between-species and within-species variation in behavior. *Current Anthropology, 44*(4), 563-571. doi:10.1086/377649
Hoppitt, W. J. E., Brown, G. R., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., & Laland, K. N. (2008). Lessons from animal teaching. *Trends in Ecology & Evolution, 23*(9), 486-493. doi:10.1016/j.tree.2008.05.008
Hoppitt, W. J. E., & Laland, K. N. (2013). *Social learning : an introduction to mechanisms, methods, and models*. Princeton: Princeton University Press.
Horstmann, A., & Hoffmann, K. P. (2005). Target selection in eye-hand coordination: Do we reach to where we look or do we look to where we reach? *Experimental Brain Research, 167*(2), 187-195. doi:10.1007/s00221-005-0038-6
Imanishi, K. (2002). *A Japanese view of nature : the world of living things*. London ; New York, NY: RoutledgeCurzon.

Ingold, T. (2000). *The perception of the environment : essays on livelihood, dwelling & skill*. London ; New York: Routledge.
Ingold, T. (2011a). *Being alive : essays on movement, knowledge and description*. London ; New York: Routledge.
Ingold, T. (2011b). Walking the Plank: Meditations on a Process of Skill *Being Alive: essays on movement, knowledge and description* (pp. 51-62). London and New York: Routledge.
Ingold, T. (2013). *Making : anthropology, archaeology, art and architecture*. London ; New York: Routledge.
Ingold, T., & Pálsson, G. s. (2013). *Biosocial becomings : integrating social and biological anthropology*. New York: Cambridge University Press.
Israel, D., & Perry, J. (1990). What Is Information. *Information, Language, and Cognition*, 1-19.
Jablonka, E. (1994). Inheritance Systems and the Evolution of New Levels of Individuality. *Journal of Theoretical Biology, 170*(3), 301-309. doi:10.1006/jtbi.1994.1191
Jablonka, E., & Lamb, M. J. (2005). *Evolution in four dimensions : genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge, Mass.: MIT Press.
Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in synthesis of proteins. *Journal of Molecular Biology, 3*(3), 318-356. doi:10.1016/s0022-2836(61)80072-7
Jacob, P. (2014). Intentionality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Winter 2014 Edition)*.
Johansson, R. S., Westling, G. R., Backstrom, A., & Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *Journal of Neuroscience, 21*(17), 6917-6932.
Juarrero, A. (1999). *Dynamics in action : intentional behavior as a complex system*. Cambridge, Mass.: MIT Press.
Kantz, H. (1994). Quantifying the Closeness of Fractal Measures. *Physical Review E, 49*(6), 5091-5097. doi:10.1103/PhysRevE.49.5091
Kauffman, S. A. (1993). *The origins of order : self-organization and selection in evolution*. New York: Oxford University Press.
Kay, L. E. (1998). A book of life? How the genome became an information system and DNA a language. *Perspectives in Biology and Medicine, 41*(4), 504-528.
Keller, E. F. (2000). *The century of the gene*. Cambridge, Mass.: Harvard University Press.
Keller, E. F. (2008). Organisms, machines, and thunderstorms: A history of self-organization, part one. *Historical Studies in the Natural Sciences, 38*(1), 45-75. doi:10.1525/hsns.2008.38.I.45
Keller, E. F. (2009). Organisms, Machines, and Thunderstorms: A History of Self-Organization, Part Two. Complexity, Emergence, and Stable Attractors. *Historical Studies in the Natural Sciences, 39*(1), 1-31. doi:10.1525/hsns.2009.39.1.1
Keller, E. F. (2010). *The mirage of a space between nature and nurture*. Durham, NC: Duke University Press.
Kellman, P. J. (2002). Perceptual learning. In C. R. Gallistel (Ed.), *Steven's Handbook of Experimental Psychology* (Vol. 3, pp. 259-299). New York: Wiley.

Kelso, J. A. S. (1995). *Dynamic patterns : the self-organization of brain and behavior*. Cambridge, Mass.: MIT Press.
Kelso, J. A. S. (2016). On the Self-Organizing Origins of Agency. *Trends Cogn Sci, 20*(7), 490-499. doi:10.1016/j.tics.2016.04.004
King, B. J. (2004). *The dynamic dance : nonvocal communication in African great apes*. Cambridge, Mass.: Harvard University Press.
Kiverstein, J. D., & Rietveld, E. (2018). Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior, Advance online publication*. doi:10.1177/1059712318772778
Klein, R. G. (2009). *The human career : human biological and cultural origins* (3rd ed.). Chicago: The University of Chicago Press.
Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-Computer Interaction, 18*(1-2), 13-49. doi:10.1207/S15327051hci1812_2
Kroeber, A. L. (1917). The Superorganic. *American Anthropologist, 19*(2), 163-213. doi:10.1525/aa.1917.19.2.02a00010
Kroeber, A. L., & Kluckhohn, C. (1952). *Culture; a critical review of concepts and definitions*. Cambridge, Mass.,: The Museum.
Kuo, Z.-Y. (1967). *The dynamics of behavior development*. New York,: Random House.
Lahlou, S. (2011a). How can we capture the subject's perspective? An evidence-based approach for the social scientist. *Social Science Information Sur Les Sciences Sociales, 50*(3-4), 607-655. doi:10.1177/0539018411411033
Lahlou, S. (2011b). *Socio-Cognitive Issues in Human-Centered Design for the Real World*.
Laland, K. N., & Galef, B. G. (2009). *The question of animal culture*. Cambridge, Mass.: Harvard University Press.
Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2001). Cultural niche construction and human evolution. *Journal of Evolutionary Biology, 14*(1), 22-33. doi:10.1046/j.1420-9101.2001.00262.x
Laland, K. N., Odling-Smee, J., & Turner, S. (2014). The role of internal and external constructive processes in evolution. *Journal of Physiology-London, 592*(11), 2413-2422. doi:10.1113/jphysiol.2014.272070
Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25-26), 3559-3565. doi:10.1016/s0042-6989(01)00102-x
Latour, B. (1993). *We have never been modern*. Cambridge, Mass.: Harvard University Press.
Lave, J., & Wenger, E. (1991). *Situated learning : legitimate peripheral participation*. Cambridge England ; New York: Cambridge University Press.
Lee, D. N. (1976). Theory of visual control of breaking based on information about time-to-collision. *Perception, 5*(4), 437-459. doi:10.1068/p050437
Lee, D. N. (1998). Guiding movement by coupling taus. *Ecological Psychology, 10*(3-4), 221-250. doi:10.1207/s15326969eco103&4_4
Lee, D. N. (2009). General Tau Theory: evolution to date. *Perception, 38*(6), 837-850. doi:10.1068/pmklee
Lee, D. N., & Reddish, P. E. (1981). Plummeting gannets - a paradigm of ecological optics. *Nature, 293*(5830), 293-294. doi:10.1038/293293a0

Lehrman, D. S. (1953). A Critique of Konrad Lorenz's Theory of Instinctive Behavior. *Quarterly Review of Biology, 28*(4), 337-363. doi:10.1086/399858
Lehrman, D. S. (1970). Semantic and conceptual issues in the nature-nurture problem. In L. R. Aronson, E. Tobach, D. S. Lehrman, & J. S. Rosenblatt (Eds.), *Development and Evolution of Behavior: Essays in Memory of T. C. Schneirla* (pp. 17–52). San Francisco: W. H. Freeman.
Lévi-Strauss, C. (1963). *Structural anthropology*. New York: Basic Books.
Lévi-Strauss, C. (1966). *The savage mind*. Chicago: University of Chicago Press.
Lévi-Strauss, C. (1992). *The view from afar* (University of Chicago Press ed.). Chicago: University of Chicago Press.
Levin, S. A. (2006). Fundamental questions in biology. *Plos Biology, 4*(9), 1471-1472. doi:10.1371/journal.pbio.0040300
Lewis, D. K. (1986). *On the plurality of worlds*. Oxford, UK ; New York, NY, USA: B. Blackwell.
Lewontin, R. C. (1983). Gene, Organism, and Environment. In D. S. Bendall (Ed.), *Evolution: From Molecules to Men* (pp. 273–285). Cambridge: Cambridge University Press.
Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior Matching in Multimodal Communication Is Synchronized. *Cognitive Science, 36*(8), 1404-1426. doi:10.1111/j.1551-6709.2012.01269.x
Markowsky, G. (2013). Information theory *Encyclopædia Britannica*: Encyclopædia Britannica, inc.
Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A, 302*(5-6), 299-307.
Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports-Review Section of Physics Letters, 438*(5-6), 237-329. doi:10.1016/j.physrep.2006.11.001
Maturana, H. R. (1980). Biology of Cognition. In H. R. Maturana & F. J. Varela (Eds.), *Autopoiesis and cognition : the realization of the living* (pp. 1-58). London: D. Reidel Pub. Co.
Maturana, H. R. (1981). Autopoiesis. In Z. M. (Ed.), *Autopoiesis: A theory of living organization.* (pp. 21–33). Boulder CO: Westview Press.
Maturana, H. R., & Varela, F. J. (1973). *De máquinas y seres vivos; una teoría sobre la organización biológia* (1. ed.). Santiago de Chile: Editorial Universitaria.
Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition : the realization of the living*. Dordrecht, Holland; Boston: D. Reidel Pub. Co.
Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge : the biological roots of human understanding* (1st ed.). Boston: New Science Library : Distributed in the United State by Random House.
Maturana, H. R., & Varela, F. J. (1992). *The tree of knowledge : the biological roots of human understanding* (Rev. ed.). Boston; New York: Shambhala.
Maynard Smith, J. (2000). The concept of information in biology. *Philosophy of Science, 67*(2), 177-194. doi:10.1086/392768
Mayr, E. (1961). Cause and Effect in Biology - Kinds of causes, predictability, and teleology are viewed by a practicing biologist. *Science, 134*(3489), 1501-1506. doi:10.1126/science.134.3489.1501
Mayr, E. (2004). *What makes biology unique? : considerations on the autonomy of a scientific discipline*. New York: Cambridge University Press.

McBrearty, S., & Brooks, A. S. (2000). The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution, 39*(5), 453-563. doi:10.1006/jhev.2000.0435

McElreath, R. (2016). *Statistical rethinking : a Bayesian course with examples in R and Stan*. Boca Raton: CRC Press/Taylor & Francis Group.

McGann, M. (2014). Enacting a social ecology: radically embodied intersubjectivity. *Frontiers in Psychology, 5*(1321). doi:10.3389/fpsyg.2014.01321

McGrew, W. C. (1992). *Chimpanzee material culture : implications for human evolution*. Cambridge England ; New York, NY, USA: Cambridge University Press.

McGrew, W. C. (2004). *The cultured chimpanzee : reflections on cultural primatology*. Cambridge, UK ; New York, NY, USA: Cambridge University Press.

McGrew, W. C., & Tutin, C. E. G. (1978). Evidence for a Social Custom in Wild Chimpanzees. *Man, 13*(2), 234-251. doi:10.2307/2800247

Mesoudi, A. (2011). *Cultural evolution : how Darwinian theory can explain human culture and synthesize the social sciences*. Chicago ; London: University of Chicago Press.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition, 66*(2), B25-B33. doi:10.1016/s0010-0277(98)00009-2

Millikan, R. G. (1984). *Language, thought, and other biological categories : new foundations for realism*. Cambridge, Mass.: MIT Press.

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of Gaze During Dynamic Scene Viewing is Predicted by Motion. *Cognitive Computation, 3*(1), 5-24. doi:10.1007/s12559-010-9074-z

Mossio, M., Montévil, M., & Longo, G. (2016). Theoretical principles for biology: Organization. *Progress in Biophysics and Molecular Biology, 122*(1), 24-35.

Nakamura, M., & Nishida, T. (2006). Subtle behavioral variation in wild chimpanzees, with special reference to Imanishi's concept of kaluchua. *Primates, 47*(1), 35-42. doi:10.1007/s10329-005-0142-z

Navarrete, A. F., Reader, S. M., Street, S. E., Whalen, A., & Laland, K. N. (2016). The coevolution of innovation and technical intelligence in primates. *Philosophical Transactions of the Royal Society B-Biological Sciences, 371*(1690). doi:10.1098/rstb.2015.0186

Neumann-Held, E. M., & Rehmann-Sutter, C. (Eds.). (2006). *Genes in development : re-reading the molecular paradigm*. Durham N.C.: Duke University Press.

Newman, S. A., & Bhat, R. (2008). Dynamical patterning modules: physico-genetic determinants of morphological development and evolution. *Physical Biology, 5*(1). doi:10.1088/1478-3975/5/1/015008

Newman, S. A., & Bhat, R. (2009). Dynamical patterning modules: a "pattern language" for development and evolution of multicellular form. *International Journal of Developmental Biology, 53*(5-6), 693-705. doi:10.1387/ijdb.072481sn

Nicolis, G., & Prigogine, I. (1977). *Self-organization in nonequilibrium systems : from dissipative structures to order through fluctuations*. New York: Wiley.

Nishida, T. (1968). The social group of wild chimpanzees in the Mahali Mountains. *Primates, 9*(3), 167-224. doi:10.1007/bf01730971

Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (1996). Niche construction. *American Naturalist, 147*(4), 641-648. doi:10.1086/285870

Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (2003). *Niche construction : the neglected process in evolution*. Princeton: Princeton University Press.
Osawa, S., Jukes, T. H., Watanabe, K., & Muto, A. (1992). Recent Evidence for Evolution of the Genetic Code. *Microbiological Reviews, 56*(1), 229-264.
Oyama, S. (1985/2000). *The ontogeny of information : developmental systems and evolution* (2nd ed.). Durham, N.C.: Duke University Press.
Oyama, S. (1999). Locating Development, Locating Developmental Systems. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 185-208). Mahwah, N.J.: Lawrence Erlbaum Associates.
Oyama, S. (2000a). Causal democracy and causal contributions in developmental systems theory. *Philosophy of Science, 67*(3), S332-S347. doi:10.1086/392830
Oyama, S. (2000b). *Evolution's eye : a systems view of the biology-culture divide*. Durham, N.C.: Duke University Press.
Oyama, S. (2009). Friends, Neighbors, and Boundaries. *Ecological Psychology, 21*(2), 147-154. doi:10.1080/10407410902877173
Oyama, S. (2015). Sustainable Development: Living with Systems. In B. Clarke (Ed.), *Earth, Life, and System: Evolution and Ecology On a Gaian Planet* (pp. 203-224). New York: Fordham University Press.
Oyama, S., Griffiths, P. E., & Gray, R. D. (2001). *Cycles of contingency : developmental systems and evolution*. Cambridge, Mass.: MIT Press.
Pagnotta, M. (2014). On the controversy over non-human culture: The reasons for disagreement and possible directions toward consensus. *Behavioural Processes, 109*, 95-100. doi:10.1016/j.beproc.2014.04.008
Papineau, D. (1987). *Reality and representation*. New York: B. Blackwell.
Péter, A. (2016). Solomon coder. Retrieved from http://solomoncoder.com/
Pitt, D. (2017). Mental Representation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition ed.).
Prigogine, I., & Nicolis, G. (1967). On Symmetry-Breaking Instabilities in Dissipative Systems. *Journal of Chemical Physics, 46*(9), 3542-3550. doi:10.1063/1.1841255
Prigogine, I., & Stengers, I. (1984). *Order out of chaos : man's new dialogue with nature* (1st ed.). Boulder, CO: New Science Library : Distributed by Random House.
R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
Rapaport, L. G., & Brown, G. R. (2008). Social influences on foraging behavior in young nonhuman primates: Learning what, where, and how to eat. *Evolutionary Anthropology, 17*(4), 189-201. doi:10.1002/evan.20180
Resende, B. D., Nagy-Reis, M. B., Lacerda, F. N., Pagnotta, M., & Savalli, C. (2014). Tufted capuchin monkeys (Sapajus sp) learning how to crack nuts: Does variability decline throughout development? *Behavioural Processes, 109*, 89-94. doi:10.1016/j.beproc.2014.09.002
Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*(6), 1045-1060. doi:10.1207/s15516709cog0000_29
Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination - Common ground and the coupling of eye movements during

dialogue. *Psychological Science, 18*(5), 407-413. doi:10.1111/j.1467-9280.2007.01914.x

Richardson, M. J., Dale, R., & Marsh, K. L. (2014). *Complex Dynamical Systems in Social and Personality Psychology Theory, Modeling, and Analysis*.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone : how culture transformed human evolution*. Chicago: University of Chicago Press.

Ridley, M. (2004). *Evolution* (3rd ed.). Malden, MA: Blackwell Pub.

Rose, S. P. R. (1981). From Causations to Translations: What Biochemists can Contribute to the Study of Behavior. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in Ethology*. Boston, MA: Springer.

Ryle, G. (1949). *The concept of mind*. London, New York,: Hutchinson's University Library.

Sarkar, S. (1996). Biological information: A skeptical look at some central dogmas of molecular biology. *Boston studies in the philosophy of science, 183*, 187-232.

Sarkar, S. (2004). Genes Encode Information for Phenotypic Traits. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science* (pp. 259-274). Malden, MA: Blackwell Publishing.

Scholz, J. P., & Kelso, J. A. S. (1990). Intentional switching between patterns of bimanual coordination depends on the intrinsic dynamics of the patterns. *Journal of Motor Behavior, 22*(1), 98-124.

Schoner, G., & Kelso, J. A. S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science, 239*(4847), 1513-1520. doi:10.1126/science.3281253

Schoner, G., Zanone, P. G., & Kelso, J. A. S. (1992). Learning as change in coordination dynamics - theory and experiment. *Journal of Motor Behavior, 24*(1), 29-48. doi:10.1080/00222895.1992.9941599

Sebanz, N., & Knoblich, G. (2009). Prediction in Joint Action: What, When, and Where. *Topics in Cognitive Science, 1*(2), 353-367. doi:10.1111/j.1756-8765.2009.01024.x

Shannon, C. E. (1948a). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*(3), 379-423.

Shannon, C. E. (1948b). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*(4), 623-656.

Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology-Human Perception and Performance, 33*(1), 201-208. doi:10.1037/0096-1523.33.1.201

Shockley, K., Butwill, M., Zbilut, J. P., & Webber, C. L. (2002). Cross recurrence quantification of coupled oscillators. *Physics Letters A, 305*(1-2), 59-69.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology-Human Perception and Performance, 29*(2), 326-332. doi:10.1037/0096-1523.29.2.326

Slagsvold, T., & Wiebe, K. L. (2007). Learning the ecological niche. *Proceedings of the Royal Society B-Biological Sciences, 274*(1606), 19-23. doi:10.1098/rspb.2006.3663

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7*(8), 343-348. doi:10.1016/S1364-6613(03)00156-6

Spencer, J. P., Perone, S., & Buss, A. T. (2011). Twenty Years and Going Strong: A Dynamic Systems Revolution in Motor and Cognitive Development. *Child Development Perspectives, 5*(4), 260-266. doi:10.1111/j.1750-8606.2011.00194.x

Stengers, I. (2000). *The invention of modern science*. Minneapolis: University of Minnesota Press.

Sterelny, K. (2000). The "genetic program" program: A commentary on Maynard Smith on information in biology. *Philosophy of Science, 67*(2), 195-201. doi:10.1086/392769

Sterelny, K., & Griffiths, P. E. (1999). *Sex and death : an introduction to philosophy of biology*. Chicago, Ill.: University of Chicago Press.

Sterelny, K., Smith, K. C., & Dickison, M. (1996). The extended replicator. *Biology & Philosophy, 11*(3), 377-403. doi:10.1007/Bf00128788

Strogatz, S. H. (2015). *Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering* (Second edition. ed.). Boulder, CO: Westview Press, a member of the Perseus Books Group.

Takemoto, C., Spremulli, L. L., Benkowski, L. A., Ueda, T., Yokogawa, T., & Watanabe, K. (2009). Unconventional decoding of the AUA codon as methionine by mitochondrial tRNAMet with the anticodon f5CAU as revealed with a mitochondrial in vitro translation system. *Nucleic Acids Research, 37*(5), 1616-1627. doi:10.1093/nar/gkp001

Thagard, P. (2014). Cognitive Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition ed.).

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, Mass.: MIT Press.

Thompson, E. (2007). *Mind in life : biology, phenomenology, and the sciences of mind*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Thompson, E., & Varela, F. J. (2001). Radical embodiment: neural dynamics and consciousness. *Trends in Cognitive Sciences, 5*(10), 418-425. doi:10.1016/S1364-6613(00)01750-2

Thornton, A., & McAuliffe, K. (2006). Teaching in wild meerkats. *Science, 313*(5784), 227-229. doi:10.1126/science.1128727

Thornton, A., & Raihani, N. J. (2008). The evolution of teaching. *Animal Behaviour, 75*, 1823-1836. doi:10.1016/j.anbehav.2007.12.014

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, Mass.: Harvard University Press.

Tomasello, M. (2003). *Constructing a language : a usage-based theory of language acquisition*. Cambridge, Mass.: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675-+. doi:10.1017/S0140525x05000129

Toren, C. (2008). An anthropology of human development: what difference does it make? *Human Development in the Twenty-First Century: Visionary Ideas from Systems Scientists*, 104-111.

Toren, C. (2012). Anthropology and Psychology. 27-41. doi:10.4135/9781446201077.n4

Turvey, M. T. (1992). Affordances and Prospective Control: An Outline of the Ontology. *Ecological Psychology, 4*(3), 173-187. doi:10.1207/s15326969eco0403_3

Turvey, M. T., & Carello, C. (2011). Obtaining information by dynamic (effortful) touching. *Philosophical Transactions: Biological Sciences, 366*(1581), 3123-3132.

Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition, 9*(3), 237-304. doi:10.1016/0010-0277(81)90002-0

Tylor, E. B. (1871). *Primitive culture: researches into the development of mythology, philosophy, religion, art, and custom*. London,: J. Murray.

van der Post, D. J., & Hogeweg, P. (2009). Cultural inheritance and diversification of diet in variable environments. *Animal Behaviour, 78*(1), 155-166. doi:10.1016/j.anbehav.2009.04.009

Van Gelder, T. (1995). What Might Cognition Be, If Not Computation. *Journal of Philosophy, 92*(7), 345-381.

van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25*(3), 785-791. doi:10.1016/j.chb.2009.02.007

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., . . . Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science, 299*(5603), 102-105. doi:10.1126/science.1078004

Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*, 229. doi:10.1038/35067550

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind : cognitive science and human experience*. Cambridge, Mass.: MIT Press.

Vivanti, G., Nadig, A., Ozonoff, S., & Rogers, S. J. (2008). What do children with autism attend to during imitation tasks? *Journal of Experimental Child Psychology, 101*(3), 186-205. doi:10.1016/j.jecp.2008.04.008

Warner, R. R. (1988). Traditionality of Mating-Site Preferences in a Coral-Reef Fish. *Nature, 335*(6192), 719-721. doi:10.1038/335719a0

Watson, J. D., & Crick, F. H. C. (1953a). Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature, 171*(4361), 964-967. doi:10.1038/171964b0

Watson, J. D., & Crick, F. H. C. (1953b). Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature, 171*(4356), 737-738. doi:10.1038/171737a0

Webber, C. L., & Zbilut, J. P. (1994). Dynamical Assessment of Physiological Systems and States Using Recurrence Plot Strategies. *Journal of Applied Physiology, 76*(2), 965-973.

Webber, C. L., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 26-94).

Whitehead, H., & Rendell, L. (2015). *The cultural lives of whales and dolphins*. Chicago: The University of Chicago Press.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., . . . Boesch, C. (1999). Cultures in chimpanzees. *Nature, 399*(6737), 682-685. doi:10.1038/21415

Whiten, A., & Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. *Advances in the Study of Behaviour*(21), 239-283.

Whiten, A., Horner, V., Litchfield, C. A., & Marshall-Pescini, S. (2004). How do apes ape? *Animal Learning & Behavior, 32*(1), 36-52. doi:10.3758/bf03196005

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE, 8*(11), e79659. doi:10.1371/journal.pone.0079659

Zbilut, J. P., Giuliani, A., & Webber, C. L. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A, 246*(1-2), 122-128. doi:10.1016/S0375-9601(98)00457-5

Zbilut, J. P., & Webber, C. L. (1992). Embeddings and Delays as Derived from Quantification of Recurrence Plots. *Physics Letters A, 171*(3-4), 199-203. doi:10.1016/0375-9601(92)90426-M

Zeleny, M. (1981). *Autopoiesis, a theory of living organizations*. New York, N.Y.: North Holland.

Appendix: Transcription of instructions (Chapter 7)

Transcriptions of the verbal instructions given by the facilitator for each task in the study reported in chapter 7. These instructions were presented as auditory stimulus for participants in conditions noVIDEO_AUDIO, noFACE_AUDIO, and FACE_AUDIO.

Star puzzle: *To complete the star puzzle, you have six identical pieces. Begin by placing two pieces side by side. Balance the second piece on top of the first. The third piece should be placed on opposite sides to create a mirror image. Now carefully lift each segment holding the three pieces together. Rotate one segment slightly so that they fit together, and you are done.*

Egg puzzle: *To complete the egg puzzle, you have two large pieces, two solid-centred pieces, four hollow-centred pieces, and a solid column. Begin by placing a large piece in the palm of your hand, and then balancing the two solid-centred pieces facing each other. Then place the second large piece on the top. Place the puzzle on its end and balance two of the hollow-centred pieces with the hollow facing up. Then turn the puzzle upside down and place the two remaining hollow-centred pieces with the hollow facing up. To complete the puzzle, insert the solid central column; and you are done.*

Barrel puzzle: *To complete the barrel puzzle, you have two large pieces, two C-shaped pieces, two E-shaped pieces, two side pieces, two top pieces, an asymmetric 'C', and a bar. To begin, place the C-shaped pieces into the centre of the large pieces and bring the two halves together. Slide one of the E-shaped pieces from the top, turn the puzzle upside down, and slide in the other E-shaped piece. You will have a large hole and a small hole. Grip the side of the small hole and pull. Position one of the side pieces and the asymmetric 'C' so that the large section is at the bottom. Place one of the top pieces so that the central bar aligns with the others. Then lock these pieces into place. Position the other side piece and turn the puzzle upside down. Enter the bar half way and then bring the bar and the remaining piece together, and you are done.*